

01 etc
 02 claim 49; Page 117; 162pp; English.
 03
 04 The present peptide is an INP60 peptide that has stimulatory and
 05 inhibitory activity on stem cell proliferation. The 4 dimensional
 06 structure of myoglobin was superimposed onto the 4 dimensional
 07 structure of the alpha chain present in human hemoglobin, and the
 08 present peptide identified as a region which is structurally different
 09 from myoglobin in 4 dimensional space. The peptide is used to
 10 inhibit stem cell proliferation, particularly during radiotherapy or
 11 chemotherapy of cancer, but more generally wherever a mammal is exposed
 12 to an agent, e.g. an antiviral, that damages or destroys stem cells. It
 13 can also be used for differential protection of normal stem cells, but
 14 not cancer cells, from chemotherapy or radiation, particularly after
 15 normal stem cells have been induced to proliferate by therapeutic
 16 treatment. The peptide can also be used for control of stem cell
 17 hypoproliferation (e.g. aplastic anaemia), for treatment or preventing
 18 stem cell exhaustion (e.g. where caused by acquired immune deficiency
 19 syndrome), and to treat or prevent immunodeficiency. INP60 can also be
 20 used for treating pain in a mammal. INP60 and related compounds
 21 reversibly inhibit or stimulate stem cells depending on the dose, so
 22 allow precise control over cycling of these cells.
 23
 24 Sequence 15 AA:
 25
 26

Query Match 100.0% Score 72; 106 bits Length 14;
 Best Local Similarity 100.0% Prod. No. 60-07;
 Matches 13; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

01 1 PPHEDLSHSAAY 14
 02 | | | | | | | | | |
 03 1 PPHEDLSHSAAY 14

RESULT 2
 AAW0010:
 ID AAW0010: standard; Peptide: 15 AA.
 AC AAW0010:
 XX

01 20 AWP 1998 (first entry)
 02
 03

04 Cytotoxic peptide 40-55 with stem cell inhibitory/stimulatory activity.
 05

06 INP60: stem cell inhibition; stimulation; proliferation; myoglobin;
 07 alpha chain; haemoglobin; radiotherapy; chemotherapy; cancer;
 08 differential protection; stem cell hypoproliferation; aplastic anaemia;
 09 stem cell exhaustion; cystic fibrosis.
 10

01 08 Hemoglobin
 02
 03

04 Key: Local/Global
 05 ID: Local/Global
 06

07 W00760922 AL:
 08
 09

10 09 OCT 1997;
 11
 12

13 03 APR 1997; 96W US05601.
 14
 15

16 03 APR 1997; 96US 0021173.
 17
 18

19 (EPON) PPHEDLSHSAAY IN:
 20
 21

22 19970001; 19970001;
 23
 24

25 Hemoglobin alpha chain deficiencies used to inhibit or stimulate
 26 stem cell proliferation, for treatment of cancer, immune depression

01 etc
 02 claim 49; Page 117; 162pp; English.
 03
 04 The present peptide is an INP60 peptide that has stimulatory and
 05 inhibitory activity on stem cell proliferation. The 4 dimensional
 06 structure of myoglobin was superimposed onto the 4 dimensional structure
 07 of the alpha chain present in human hemoglobin, and the present peptide
 08 identified as a region which is structurally different from myoglobin in
 09 4 dimensional space. The peptide was cyclized to more closely
 10 approximate the loop found in the native alpha chain. The 2 end cysteine
 11 residues can also be joined by a carbon bridge. The peptide is used to
 12 inhibit stem cell proliferation, particularly during radiotherapy or
 13 chemotherapy of cancer, but more generally wherever a mammal is exposed
 14 to an agent, e.g. an antiviral, that damages or destroys stem cells. It
 15 can also be used for differential protection of normal stem cells, but
 16 not cancer cells, from chemotherapy or radiation, particularly after
 17 normal stem cells have been induced to proliferate by therapeutic
 18 treatment. The peptide can also be used for control of stem cell
 19 hypoproliferation (e.g. aplastic anaemia), for treatment or preventing
 20 stem cell exhaustion (e.g. where caused by acquired immune deficiency
 21 syndrome), and to treat or prevent immunodeficiency. INP60 can also be
 22 used for treating pain in a mammal. INP60 and related compounds
 23 reversibly inhibit or stimulate stem cells depending on the dose, so
 24 allow precise control over cycling of these cells.
 25
 26 Sequence 15 AA:
 27
 28

Query Match 100.0% Score 72; 106 bits Length 15;
 Best Local Similarity 100.0% Prod. No. 60-07;
 Matches 13; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

01 1 PPHEDLSHSAAY 14
 02 | | | | | | | | | |
 03 2 PPHEDLSHSAAY 14

RESULT 3
 AAW1984:
 ID AAW1984: standard; Peptide: 16 AA.
 AC AAW1984:
 XX

01 09 DEC 1999 (first entry)
 02
 03

04 Rheumatoid arthritis diagnostic protein isoform peptide #134.
 05

06 Human; rheumatoid arthritis; RA; diagnosis; RFL; RAGE; detection;
 07 Rheumatoid arthritis diagnostic reference; RFL; regional fluid;
 08 Rheumatoid arthritis diagnostic protein isoform screening;
 09 expression reference protein isoform; prognosis.
 10

11 08 Homo sapiens
 12
 13

14 W0947925 AL:
 15
 16

17 23 SEP 1999;
 18
 19

20 15-MAR-1999; 99W0-G000763.
 21
 22

23 13 MAR 1999; 98GB 0005477.
 24
 25

26 (OXF) 1 OXFOKD GYVOSTENES DR LTH.
 27
 28

29 Parokh Bv, Patel TP, Townsend PK;
 30
 31

32 WPI: 1999-071871/48.
 33
 34

35 diagnosis of human rheumatoid arthritis by two dimensional
 36 electrophoresis -
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

XX A method has been developed for the diagnosis of human rheumatoid arthritis (RA) using two-dimensional electrophoresis to generate a two-dimensional array of features. The method can be used for screening, diagnosis and prognosis of RA in a subject or for monitoring the effect of an anti-RA drug or therapy administered to a subject. The method comprises: (a) analysing a sample of serum or plasma and optionally synovial fluid by two-dimensional electrophoresis, to generate a two-dimensional array of features; (b) identifying at least one chosen feature whose relative abundance correlates with the presence or absence of RA; and (c) comparing the abundance of each chosen feature in the sample with the abundance of that chosen feature in serum or plasma from one or more persons without RA, where the relative abundance of the chosen feature or features in the sample indicates the presence or absence of RA in the subject. The method can also be used in clinical studies for testing drugs for therapy of RA, for purification of RA diagnostic protein isoforms (RPIs), and for production of antibodies to RPIs. The RA diagnostic feature (WAF) proteins can be used to identify compounds that promote or inhibit their activity, which are then used as RA drugs. Nucleic acid encoding RPIs can be used in gene therapy protocols. AA441844 to AA442100 represent RPI peptides. AA442101 to AA442106 represent expression-informed protein isoform peptides and AA442106 to AA442608 represent peptide fragments for RPIs, which are all used in the exemplification of the present invention.

Sequence: 16 AA:

Query Match	100.0%;	Score 72;	DB 20;	Length 16;
Best Local Similarity	100.0%;	Pred. No. 5.4e-07;		
Matches	13;	Conservative	0;	Indels
		Mismatches	0;	Gaps
				0;

```
QY 1 PPHPLSHOSAGV 13
    ||| |||||
Db 3 fPpdlshsagv 15
```

RESULT 4

TD AAW66275 standard; peptide; 19 AA.

AC: AAW66275;

DT 23-NOV-1998 (first entry)

Dr. Circularly permuted di-alpha haemaglobin-like peptide fragment.

KW haemoglobin; circularly permuted; N-terminus; C-terminus; multimeric
KW permuted haemoglobin; gobin; pibo; red blood cell; blood transfusion;
KW sickle cell; stroke.

OS Synthetic

PN W09838211-A2

PD 03-SEF-1998.

PI 27-FRH-1948; 98WO-US03836.

PR 28-FEB-1997; 97US-0039657.

PA (SOMA-); SOMATOTROPIN INC.

PI best HA, 'olins I';

DR WP I; 1956-481141/41.
XX

PT New hemooglobin ferrihem with non-naturally occurring termini -
 PT useful for the treatment of any condition where therapy includes red
 blood cells or oxygen delivery
 XX
 Disclosure date: 10/27/2011; Ertlisch.

The invention relates to a haemoglobin permutoxin (HPH), containing at least one substitution with non-naturally occurring N and C terminal. Also claimed is a nucleic acid encoding the above HPH. pHPb can be used as a substitute for red blood cells whenever red blood cells or oxygen solubility is required, particularly in the treatment of haemorrhages, traumas, and surgeries requiring blood transfusion or perfusion of the patient's own blood and anaemia, to deliver oxygen to tissue sites that red blood cells cannot penetrate, including site downstream of an obstruction as in thrombi, stroke cell occlusions, arterial occlusions, angioplasty balloons, and surgical instrumentations, and to treat tissue ischaemias including stroke, emerging stroke, transient ischaemic attacks, myocardial stunning and fibrillation, acute or unstable angina, and to treat a heart. The invention can also be used to deliver drugs and for *in vivo* imaging. The permutoxins have increased stability and expression during E. coli synthesis and a greater *in vivo* circulating half-life compared to the fused Hb molecules of prior art. The present sequence is shown in the specification.

SQ sequence 19 AA;

Query Match 100.0%; Score 72; PR 19; Length 19;
 Best Local Similarity 100.0%; Prev. No. 9.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

QY 1 FPHFDLSHGSAUY 13
    | | | | | | | | | |
Db 6 fphfdlshgsayv 18

```

RESULTS

1) AAR29606 standard; Protein; 141 AA

AC AAF29606;

DI 15-JAN-1993 (first entry)

1D: Mutant human alpha-globin 1.

KW Tetraepot intramolecular crosslink; oxygen exchange; erythrocytes; HIV

05 Synthetic

[illegible]

/*tāq= a

XX
XX
W00011303-2XX
XX
06-III-1603

XX	0150-11306634
DE	20-1332-1001

XX 70-FWC-1990, 90ME-0820825
PRXX
PA CIVIL -) UNIVERSITY OF ALABAMA PRESS (JUN 1971)

XX
XX
PI McJuno SI Jones TM

XX
XX
WPT - 1643-250626.231)

XX N-PSIB; AAQ26424 .
XX

PT Transgenic, cross-linked polymeric human haemoglobin - used as a cell-free blood substitute

CC This sequence was not disclosed in the specification. The alpha
CC alpha gene sequence was obtained from Swiss Pat. no. AA901922. The
CC mutation was then added to the sequence. The sequence was mutated by
CC polymerase chain reaction. The mutation introduced produced a

1	26	FEH 19961	6003 0160766	
2	26	FEH 19961	6003 0160766	
3	26	FEH 19961	6003 0160766	
4	15	100 1000	6003 0160729	
5	17	FEH 19961	6003 0160766	
6	05	MAE 19961	6003 0160766	
7	08	100 1000	6003 0160766	
8	27	001 19961	6003 0160729	
9	15	100 1000	6003 0160729	
10	26	100 1000	6003 0160766	
11	01	MAE 19961	6003 0160766	
12	08	100 1000	6003 0160766	
13	27	001 19961	6003 0160729	
14	15	100 1000	6003 0160729	
15	26	100 1000	6003 0160766	
16	01	MAE 19961	6003 0160766	
17	08	100 1000	6003 0160766	
18	27	001 19961	6003 0160729	
19	15	100 1000	6003 0160729	
20	26	100 1000	6003 0160766	
21	01	MAE 19961	6003 0160766	
22	08	100 1000	6003 0160766	
23	27	001 19961	6003 0160729	
24	15	100 1000	6003 0160729	
25	26	100 1000	6003 0160766	
26	01	MAE 19961	6003 0160766	
27	08	100 1000	6003 0160766	
28	27	001 19961	6003 0160729	
29	15	100 1000	6003 0160729	
30	26	100 1000	6003 0160766	
31	01	MAE 19961	6003 0160766	
32	08	100 1000	6003 0160766	
33	27	001 19961	6003 0160729	
34	15	100 1000	6003 0160729	
35	26	100 1000	6003 0160766	
36	01	MAE 19961	6003 0160766	
37	08	100 1000	6003 0160766	
38	27	001 19961	6003 0160729	
39	15	100 1000	6003 0160729	
40	26	100 1000	6003 0160766	
41	01	MAE 19961	6003 0160766	
42	08	100 1000	6003 0160766	
43	27	001 19961	6003 0160729	
44	15	100 1000	6003 0160729	
45	26	100 1000	6003 0160766	
46	01	MAE 19961	6003 0160766	
47	08	100 1000	6003 0160766	
48	27	001 19961	6003 0160729	
49	15	100 1000	6003 0160729	
50	26	100 1000	6003 0160766	
51	01	MAE 19961	6003 0160766	
52	08	100 1000	6003 0160766	
53	27	001 19961	6003 0160729	
54	15	100 1000	6003 0160729	
55	26	100 1000	6003 0160766	
56	01	MAE 19961	6003 0160766	
57	08	100 1000	6003 0160766	
58	27	001 19961	6003 0160729	
59	15	100 1000	6003 0160729	
60	26	100 1000	6003 0160766	
61	01	MAE 19961	6003 0160766	
62	08	100 1000	6003 0160766	
63	27	001 19961	6003 0160729	
64	15	100 1000	6003 0160729	
65	26	100 1000	6003 0160766	
66	01	MAE 19961	6003 0160766	
67	08	100 1000	6003 0160766	
68	27	001 19961	6003 0160729	
69	15	100 1000	6003 0160729	
70	26	100 1000	6003 0160766	
71	01	MAE 19961	6003 0160766	
72	08	100 1000	6003 0160766	
73	27	001 19961	6003 0160729	
74	15	100 1000	6003 0160729	
75	26	100 1000	6003 0160766	
76	01	MAE 19961	6003 0160766	
77	08	100 1000	6003 0160766	
78	27	001 19961	6003 0160729	
79	15	100 1000	6003 0160729	
80	26	100 1000	6003 0160766	
81	01	MAE 19961	6003 0160766	
82	08	100 1000	6003 0160766	
83	27	001 19961	6003 0160729	
84	15	100 1000	6003 0160729	
85	26	100 1000	6003 0160766	
86	01	MAE 19961	6003 0160766	
87	08	100 1000	6003 0160766	
88	27	001 19961	6003 0160729	
89	15	100 1000	6003 0160729	
90	26	100 1000	6003 0160766	
91	01	MAE 19961	6003 0160766	
92	08	100 1000	6003 0160766	
93	27	001 19961	6003 0160729	
94	15	100 1000	6003 0160729	
95	26	100 1000	6003 0160766	
96	01	MAE 19961	6003 0160766	
97	08	100 1000	6003 0160766	
98	27	001 19961	6003 0160729	
99	15	100 1000	6003 0160729	
100	26	100 1000	6003 0160766	

[illegible]

PR 12-JUN-1992: 92US-0897648.
 PR 08-DEC-1992: 92US-0987890.
 PR 15-MAR-1994: 94US-0030897.
 XX
 XX
 XX (DNXD-) DNK CORP.
 PI Holzman SH, Kellor H, Kumar P, Logan JS, O'Connell JK,
 PI Parsons CT, Pilder SH, Pinkert CA, Sharma A, Swanson ME;
 PI White SP;
 PR 1994-007675/91.
 XX
 XX
 XX Transgenic pigs which produce human haemoglobin: conty. DNA encoding
 PI human alpha globin and human beta globin linked to promoters
 PS
 PS Claim 14: 117pp: English.
 XX
 XX The sequences given in AAR44518-20 represent mutant versions of the
 CC adult human alpha globin sequence. These sequences were not given in
 CC the specification but were generated using the known alpha globin
 CC sequence. These sequences were used in the production of constructs which
 CC were used to produce the transgenic pigs of the invention. These
 CC constructs contain the pig beta globin promoter sequence operably linked
 CC to the human alpha or beta globin genes. Pigs containing these
 CC constructs express human haemoglobin (Hb) in their erythrocytes and are
 CC healthy, suffering no deleterious side effects as a result of
 CC heterologous Hb production. They can be used as an efficient and
 CC economical source of human Hb that can be used for transfusion and other
 CC medical applications. Also, a human/pig hybrid Hb can be produced which
 CC exhibits a p50 that is higher than that of native human or pig Hb.
 XX
 XX Sequence 141 AA.
 SU
 Query Match: 100.0%; Score 72; DB 15; Length 141;
 Best Local Similarity 100.0%; Pred. No. 6, 40-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPHFLLSHSGAVY 13
 DB 43 fphdshsgav 55
 RESULT 13
 AAR44518
 ID AAR44518 standard; Protein: 141 AA.
 XX
 AC AAR44518;
 XX
 AC
 XX 23-JUN-1994 (first entry)
 XX
 DE Adult human alpha-globin 1106p
 XX
 KW Adult; pig; beta; globin; gene; regulatory region; promoter;
 KW transgenic pig; human; alpha; haemoglobin; Hb; erythrocyte;
 KW transfection; human/pig; hybrid Hb
 XX
 OS Homo sapiens.
 XX
 OS
 FH Key Location/Qualifiers
 FT Misc difference 106
 FT /Label - L106F
 XX
 PN W09325071-A.
 PN
 XX 23-DEC-1994.
 PD
 XX 11-JUN-1994: 94US-0030897.
 PR 12-JUN-1992: 92US-0897648.
 PR 08-DEC-1992: 92US-0987890.
 PR 15-MAR-1994: 94US-0030897.
 XX
 XX

PA (DNXD-) DNK CORP.
 XX
 XX Holzman SH, Kellor H, Kumar P, Logan JS, O'Connell JK,
 PI Parsons CT, Pilder SH, Pinkert CA, Sharma A, Swanson ME;
 PI White SP;
 PR 1994-007675/91
 XX
 XX
 XX Transgenic pigs which produce human haemoglobin: conty. DNA encoding
 PI human alpha globin and human beta globin linked to promoters
 PS
 PS Claim 14: 117pp: English.
 XX
 XX The sequences given in AAR44518-20 represent mutant versions of the
 CC adult human alpha globin sequence. These sequences were not given in
 CC the specification but were generated using the known alpha globin
 CC sequence. These sequences were used in the production of constructs which
 CC were used to produce the transgenic pigs of the invention. These
 CC constructs contain the pig beta globin promoter sequence operably linked
 CC to the human alpha or beta globin genes. Pigs containing these
 CC constructs express human haemoglobin (Hb) in their erythrocytes and are
 CC healthy, suffering no deleterious side effects as a result of
 CC heterologous Hb production. They can be used as an efficient and
 CC economical source of human Hb that can be used for transfusion and other
 CC medical applications. Also, a human/pig hybrid Hb can be produced which
 CC exhibits a p50 that is higher than that of native human or pig Hb.
 XX
 XX Sequence 141 AA:
 SU
 Query Match: 100.0%; Score 72; DB 15; Length 141;
 Best Local Similarity 100.0%; Pred. No. 6, 40-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPHFLLSHSGAVY 13
 DB 43 fphdshsgav 55
 RESULT 14
 AAR44519
 ID AAR44519 standard; Protein: 141 AA.
 XX
 AC AAR44519;
 XX
 AC
 XX 23-JUN-1994 (first entry)
 XX
 DE Adult human alpha-globin V107S.
 XX
 KW Adult; pig; beta; globin; gene; regulatory region; promoter;
 KW transgenic pig; human; alpha; haemoglobin; Hb; erythrocyte;
 KW transfection; human/pig; hybrid Hb.
 XX
 OS Homo sapiens.
 XX
 OS
 FH Key Location/Qualifiers
 FT Misc difference 107
 FT /Label - V107S
 XX
 PN W09325071-A
 PN
 XX 23-DEC-1994.
 PD
 XX 11-JUN-1994: 94US-0030897.
 PR 12-JUN-1992: 92US-0897648.
 PR 08-DEC-1992: 92US-0987890.
 PR 15-MAR-1994: 94US-0030897.
 XX
 XX
 XX (DNXD-) DNK CORP.
 PI Holzman SH, Kellor H, Kumar P, Logan JS, O'Connell JK,
 PI Parsons CT, Pilder SH, Pinkert CA, Sharma A, Swanson ME;
 XX

TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-627-173-2

Query Match: 100.0%; Score 72; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 4,8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PFHFDLSHGSAQV 13
DB 2 PFHFDLSHGSAQV 14

RESULT 5
US-08-535-882A-2
Sequence 2: Application US/08/535882A
Patent No. 5949391
GENERAL INFORMATION:
APPLICANT: TSKLOVA, IRENA
APPLICANT: WOLFE, STEPHEN D.
TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH CLIFRE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPIER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/02/535,882A
FILING DATE: 28-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1331-177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-535-882A-2

Query Match: 100.0%; Score 72; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 4,8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PFHFDLSHGSAQV 13
DB 2 PFHFDLSHGSAQV 14

RESULT 6
US-09-005-546-2
Sequence 2: Application US/09/005546
Patent No. 6090782
GENERAL INFORMATION:
APPLICANT: TSKLOVA, IRENA
APPLICANT: WOLFE, STEPHEN D.
TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH CLIFRE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201 4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPIER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,546-2
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/535,882
FILING DATE: 28-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1331-177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-005-546-2

Query Match: 100.0%; Score 72; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 4,8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PFHFDLSHGSAQV 13
DB 2 PFHFDLSHGSAQV 14

RESULT 7
US-07-923-692C-8
Sequence 8: Application US/07/923692C
Patent No. 5316931
GENERAL INFORMATION:
APPLICANT: Ionsen, Jon
APPLICANT: Ionsen, William O.
APPLICANT: Grantham, George L.
APPLICANT: Turpen, Thomas H.
APPLICANT: Turpen, Ann Myers
APPLICANT: Garger, Stephen J.
APPLICANT: Griffin, Laurence K.
TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Limbach & Limbach
STREET: 2001 Ferry Building
CITY: San Francisco

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 641,617
 FILING DATE: 16-JAN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 310,881
 FILING DATE: 17-FEB-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 160,766
 FILING DATE: 26-FEB-1988
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 160,771
 FILING DATE: 26-FEB-1988
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 347,637
 FILING DATE: 05-MAY-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 363,138
 FILING DATE: 08-JUN-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 219,279
 FILING DATE: 19-JUL-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Ballwin, Albert P.
 REGISTRATION NUMBER: 28,957
 REFERENCE/AGENT NUMBER: BIOG-20121 USA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-433-4150
 TELEFAX: 415-433-8716
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLTYPE: protein
 US-08-184-237-8

Query Match 100.0% Score 72, DB 1, Length 141,
 Best Local Similarity 100.0% Pctid. NO. 50-05;
 Matches 13, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 PFPHLSHSAVY 13
 DB 43 PFPHLSHSAVY 55

RESULT 10
 US-08-240-712-17
 Sequence 7, Application US/08240712
 Patent No. 5599907
 GENERAL INFORMATION:
 APPLICANT: ANDERSON, DAVID C.
 APPLICANT: MATHEWS, ANTONY JAMES
 APPLICANT: STELLER, GARY L.
 TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
 TITLE OF INVENTION: HEMOGLOBINS
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Broadway and Melmark
 STREET: 419 Seventh Street, N.W., Suite 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: us-08-240-712
 FILING DATE: 05-MAY-1994
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08240712
 FILING DATE: 13-MAY-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: COOPER, IVER P.
 REGISTRATION NUMBER: 28,005
 REFERENCE/AGENT NUMBER: ANDERSON 6
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 TELEX: 248633
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLTYPE: protein
 US-08-240-712-17

Query Match 100.0% Score 72, DB 1, Length 141,
 Best Local Similarity 100.0% Pctid. NO. 60-06;
 Matches 13, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 PFPHLSHSAVY 13
 DB 43 PFPHLSHSAVY 55

RESULT 11
 US-08-240-712-26
 Sequence 26, Application US/08240712
 Patent No. 5599907
 GENERAL INFORMATION:
 APPLICANT: ANDERSON, DAVID C.
 APPLICANT: MATHEWS, ANTONY JAMES
 APPLICANT: STELLER, GARY L.
 TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
 TITLE OF INVENTION: HEMOGLOBINS
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Broadway and Melmark
 STREET: 419 Seventh Street, N.W., Suite 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: us-08-240-712
 FILING DATE: 05-MAY-1994
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: us-08-240-712
 FILING DATE: 13-MAY-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: COOPER, IVER P.
 REGISTRATION NUMBER: 28,005
 REFERENCE/AGENT NUMBER: ANDERSON-6
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 TELEX: 248633
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: single

SOURCE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 05/094443, 890
 FILING DATE: 31-MAY-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/240,712
 FILING DATE: 09-MAY-1994
 APPLICATION NUMBER: PCT/0592/09752
 FILING DATE: 14-MAY-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: COOPER, IVER P
 REGISTRATION NUMBER: 28,005
 REFERENCE/DCKET NUMBER: ANDERSON-6
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 TELEX: 248633
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-443-890-17

Query Match: 100.0%; Score 72; DB 1; Length 141;
 Best Local Similarity: 100.0%; Pred. No. 60-06;
 Matches: 13; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 1 PFPHDSHSAV 13
 DB 43 PFPHDSHSAV 55

RESULT 15
 US-08-443-890-26
 Sequence 26, Application US/08443890
 Patent No. 5739011
 GENERAL INFORMATION:
 APPLICANT: ANDERSON, DAVID C.
 APPLICANT: MAHEWS, ANTONY JAMES
 APPLICANT: STETLER, GARY L.
 TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Broadway and Neimark
 STREET: 419 Seventh Street, N.W., Suite 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/443,890
 FILING DATE: 31-MAY-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/240,712
 FILING DATE: 09-MAY-1994
 APPLICATION NUMBER: PCT/0592/09752
 FILING DATE: 14-MAY-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: COOPER, IVER P
 REGISTRATION NUMBER: 28,005
 REFERENCE/DCKET NUMBER: ANDERSON-6

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 TELEX: 248633
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-443-890-26

Query Match: 100.0%; Score 72; DB 1; Length 141;
 Best Local Similarity: 100.0%; Pred. No. 60-06;
 Matches: 13; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 1 PFPHDSHSAV 13
 DB 43 PFPHDSHSAV 55

Search completed: Jun 21, 2001, 10:25:09
 Job Time: 258 sec



GenCore version 4.5
Copyright (c) 1998 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 21, 2001, 19:24:01 Seq. # 16, 8, 0 words

Title:	US-08-832-443E-1
Port of origin:	72

Sequence: 1 PPHTDISHGSAQY 13

Sequence: 1 PPHTDISHGSAQY 13

Scoring table: `HL0SUM62`

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 21924

Maximum DB seq length: 2000000000

Maximum DB seq length: 2000000000

Post-processing: Minimum Match (98)

Listing first 45 summaries

```
Database :          plr_68:*
1: plr1:*
2: plr2:*
3: plr3:*
4: plr4:*
```

Pred. R^2 is the number of positives predicted by eqn (1) to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match	Length	DB	ID	Description
1	72	100.0	140	1	HABRKA	hemoglobin alpha c
2	72	100.0	141	1	HACZLP	hemoglobin alpha c
3	72	100.0	141	1	HACVY	hemoglobin alpha c
4	72	100.0	141	1	HAMQIP	hemoglobin alpha c
5	72	100.0	141	1	HAMQJ	hemoglobin alpha c
6	72	100.0	141	1	HAMQC	hemoglobin alpha c
7	72	100.0	141	1	HAMQIP	hemoglobin alpha c
8	72	100.0	141	1	HAMQIP	hemoglobin alpha c
9	72	100.0	141	1	HAMKP	hemoglobin alpha c
10	72	100.0	141	1	HACICB	hemoglobin alpha c
11	72	100.0	141	1	HAMCA	hemoglobin alpha c
12	72	100.0	141	1	HAMQF	hemoglobin alpha c
13	72	100.0	141	1	HAMOB	hemoglobin alpha c
14	72	100.0	141	1	HAIEN	hemoglobin alpha c
15	72	100.0	141	1	HAIKS	hemoglobin alpha c
16	72	100.0	141	1	HAGC	hemoglobin alpha c
17	72	100.0	141	1	HAOEE	hemoglobin alpha c
18	72	100.0	141	1	HAPID	hemoglobin alpha c
19	72	100.0	141	1	HAIJDA	hemoglobin alpha c
20	72	100.0	141	1	HAIJE	hemoglobin alpha c
21	72	100.0	141	1	HATX	hemoglobin alpha c
22	72	100.0	141	1	HAIQIP	hemoglobin alpha c
23	72	100.0	141	1	HAPID	hemoglobin alpha c
24	72	100.0	141	1	HAKES	hemoglobin alpha c
25	72	100.0	141	1	HAKTE	hemoglobin alpha c
26	72	100.0	141	1	HAMHIF	hemoglobin alpha c
27	72	100.0	141	1	HAKVLE	hemoglobin alpha c
28	72	100.0	141	1	HAKNIF	hemoglobin alpha c
29	72	100.0	141	1	HAKNIF	hemoglobin alpha c

30	72	100.0	1.41	1	HAGK2E	hemoglobin alpha-1
31	72	100.0	1.41	1	HAGN2E	hemoglobin alpha-1
32	72	100.0	1.41	1	HABF	hemoglobin alpha-1
33	72	100.0	1.41	1	HABV	hemoglobin alpha-1
34	72	100.0	1.41	1	HABD	hemoglobin alpha-1
35	72	100.0	1.41	1	HABOK	hemoglobin alpha-1
36	72	100.0	1.41	1	HABOZ	hemoglobin alpha-1
37	72	100.0	1.41	1	HABOY	hemoglobin alpha-1
38	72	100.0	1.41	1	HABPI	hemoglobin alpha-1
39	72	100.0	1.41	1	HABP2E	hemoglobin alpha-1
40	72	100.0	1.41	1	HABNM	hemoglobin alpha-1
41	72	100.0	1.41	1	HABE	hemoglobin alpha-1
42	72	100.0	1.41	1	HABV2	hemoglobin alpha-1
43	72	100.0	1.41	1	HABOG	hemoglobin alpha-1
44	72	100.0	1.41	1	HABGT	hemoglobin alpha-1
45	72	100.0	1.41	2	S20278	hemoglobin alpha-1

ALUMINUM

```

RESULT 1
HABOKA
hemoglobin alpha chain - greater kudu
C:Species: Tragelaphus streptoceros (greater kudu)
C:Date: 04-100-1086 #sequence_revision 04-100-1086 #seq_change 03-Mar-2000
C:Accession: A02293
R:Rohrwald, K.; Wiesner, H.; Brantlitz, G.
BIOL. Chem. Hoppe-Seyler 366, 495-502, 1985
A>Title: Primary structure of the hemoglobins from the greater kudu antelope (Tragelaphus streptoceros)
A:Reference number: A90684 MDID:85279893
A:Accession: A02293
A:Molecule type: protein
A:Residues: 1-140 <RCD>
C:Superfamily: globin
C:Keywords: blood, chromatocytin, erythrocyte, hemoglobin, hemoglobin, oxygen carrier, heme, heme iron (heme) (proximal axial ligand) #status predicted
P:77-binding site, heme iron (heme) (proximal axial ligand) #status predicted
P:68-binding site, heme iron (heme) (proximal axial ligand) #status predicted

Query Match
Best local Similarity 100.0%: Score 72: DB 1: Length 140:
Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

CY 1 PPHDUSHGSAOV 13
+++++
DB 42 PPHDUSHGSAOV 54

RESULT 2
HACZP
hemoglobin alpha chain - pygmy chimpanzee
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #seq_change 03-Mar-2000
C:Accession: C93303
R:Goodman, M.; Brantlitz, G.; Stangl, A.; Schrank, B.
Nature 303, 546-548, 1983
A>Title: Evidence on human origins from haemoglobins of African apes.
A:Reference number: A93303 MDID:84219265
A:Accession: C93303
A:Molecule type: protein
A:Residues: 1-141 <CD>
C:Superfamily: globin
C:Keywords: blood, chromatocytin, erythrocyte, hemoglobin, hemoglobin, oxygen carrier, heme, heme iron (heme) (proximal axial ligand) #status predicted
P:58-binding site, heme iron (heme) (proximal axial ligand) #status predicted
P:67-binding site, heme iron (heme) (proximal axial ligand) #status predicted

```


R:Mahoney, W.C.; Nute, P.E.
Arch. Biochem. Biophys. 196, 64-72, 1979
A:Title: Hemoglobin alpha-chain variation in macaques: primary structures of the alpha(I)
A:Reference number: S06514; MUID:80063760
A:Accession: S06514
A:Molecule type: Protein
A:Residues: 1-141 <MA>
C:Superfamily: globin; globin homology
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein
F:2-141/Domain: globin homology <GLB>
F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 8, 2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLSHGSACV 13
DB 43 FPHFDLSHGSACV 55

RESULT 8
HAMQZP
hemoglobin alpha II chain - pig-tailed macaque
C:Species: Macaca nemestrina (pig-tailed macaque)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 03-Mar-2000
C:Accession: S07670
R:Mahoney, W.C.; Nute, P.E.
Arch. Biochem. Biophys. 196, 64-72, 1979
A:Title: Hemoglobin alpha-chain variation in macaques: primary structures of the alpha(I)
A:Reference number: S06514; MUID:80063760
A:Accession: S07670
A:Molecule type: protein
A:Residues: 1-141 <MA>
C:Superfamily: globin; globin homology
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein
F:2-141/Domain: globin homology <GLB>
F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 8, 2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLSHGSACV 13
DB 43 FPHFDLSHGSACV 55

RESULT 9
HAMKP
hemoglobin alpha chain - black-handed spider monkey
C:Species: Atelos geoffroyi (black-handed spider monkey)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 03-Mar-2000
C:Accession: A02253
R:Maranda, G.; Maita, T.; Sugiyama, Y.; Hayashi, M.; Ota, Y.; Ariga, A.; Goodman, M.
Hoppe Seyler's Z. Physiol. Chem. 354, 1513-1516, 1973
A:Title: Studies on the primary structures of alpha and beta polypeptide chains of adult
A:Reference number: A02253; MUID:75059125
A:Accession: A02253
A:Molecule type: protein
A:Residues: 1-141 <MA>
C:Superfamily: globin; globin homology
C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier
F:2-141/Domain: globin homology <GLB>
F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 8, 2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLSHGSACV 13
DB 43 FPHFDLSHGSACV 55

Best Local Similarity 100.0%; Pred. No. 8, 2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLSHGSACV 13
DB 43 FPHFDLSHGSACV 55

RESULT 10
HACCB
hemoglobin alpha chain - black-tailed marmoset
C:Species: Callithrix jacchus (black-tailed marmoset)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 03-Mar-2000
C:Accession: S06512
R:Maita, T.; Hayashida, M.; Matsuda, G.
J. Biochem. 95, 805-813, 1984
A:Title: Primary structures of adult hemoglobins of silver marmoset, Callithrix jacchus
A:Reference number: S06512; MUID:84212383
A:Accession: S06512
A:Molecule type: protein
A:Residues: 1-141 <MA>
C:Superfamily: globin; globin homology
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein
F:2-141/Domain: globin homology <GLB>
F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 8, 2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLSHGSACV 13
DB 43 FPHFDLSHGSACV 55

RESULT 11
HAMOA
hemoglobin alpha chain - brown capuchin
C:Species: Cebus apella (brown capuchin, black capped capuchin)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 03-Mar-2000
C:Accession: A02254
R:Maranda, G.; Maita, T.; Watanabe, F.; Ariga, A.; Goodman, M.; Hoppe Seyler's Z. Physiol. Chem. 354, 1513-1516, 1973
A:Title: The amino acid sequences of the alpha and beta polypeptide chains of adult h
A:Reference number: A02254; MUID:75059125
A:Accession: A02254
A:Molecule type: protein
A:Residues: 1-141 <MA>
C:Superfamily: globin; globin homology
C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier
F:2-141/Domain: globin homology <GLB>
F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 8, 2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDLSHGSACV 13
DB 43 FPHFDLSHGSACV 55

RESULT 12
HAMOF
hemoglobin alpha chain - brown-headed tamarin
C:Species: Saguinus fuscicollis (brown-headed tamarin, saddle-backed tamarin)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 03-Mar-2000
C:Accession: A02255
R:Llin, E.D.; Kim, Y.K.; Chernoff, A.L.

hemoglobin: 24, 47, 48, 1976

Active: Primary structure of the hemoglobin (subunits 1, 2, 3, 4) hemoglobin: 1, 1980

Accession number: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

Accession: A02260

GeneCore version 4.5
Copyright (c) 1993 - 2000 Compton Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 10:25:11 : Search time 11.96 Seconds
(without alignments)

37,234 Million cell updates/sec

Title: US-08-832-443b-1

Perfect score: 72

Sequence: 1 PFHFDLSHSAQV 13

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39:*

Result No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	100.0	140	1 HBA_TFAST	P04237 Tragalaphus
2	72	100.0	141	1 HBA_TADAR	P11755 Tadarida br
3	72	100.0	141	1 HBA2_ARC3A	P11330 arctoccephal
4	72	100.0	141	1 HBA2_HUSM1	P01968 bos mutus g
5	72	100.0	141	1 HBA_ARC3A	P01971 alces alces
6	72	100.0	141	1 HBA_ARC3A	P14387 antilocapra
7	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
8	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
9	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
10	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
11	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
12	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
13	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
14	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
15	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
16	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
17	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
18	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
19	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
20	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
21	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
22	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
23	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
24	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
25	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
26	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
27	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
28	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
29	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
30	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
31	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
32	72	100.0	141	1 HBA_ARC3A	P01927 alces alces
33	72	100.0	141	1 HBA_ARC3A	P01927 alces alces

34	72	100.0	141	1 HBA_MACAS	P21766 macaca assa
35	72	100.0	141	1 HBA_MACAS	P21767 macaca assa
36	72	100.0	141	1 HBA_MACAS	P21768 macaca assa
37	72	100.0	141	1 HBA_MACAS	P21769 macaca assa
38	72	100.0	141	1 HBA_MACAS	P21770 macaca assa
39	72	100.0	141	1 HBA_MACAS	P21771 macaca assa
40	72	100.0	141	1 HBA_MACAS	P21772 macaca assa
41	72	100.0	141	1 HBA_MACAS	P21773 macaca assa
42	72	100.0	141	1 HBA_MACAS	P21774 macaca assa
43	72	100.0	141	1 HBA_MACAS	P21775 macaca assa
44	72	100.0	141	1 HBA_MACAS	P21776 macaca assa
45	72	100.0	141	1 HBA_MACAS	P21777 macaca assa

ALIGNMENTS

RESULT 1	HBA_TFAST	STANDARD	PRT: 140 AA.
ID	HBA_TFAST		
AC	P04237		
DT	20-MAR-1987 (Rel. 04, Created)		
DT	20-MAR-1987 (Rel. 04, Last sequence update)		
DT	01-JUN-1994 (Rel. 29, Last annotation update)		
DE	HEMORRHOIDIN ALPHA CHAIN.		
OS	Tragalaphus streptoceros (Greater kudu).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eulalia; Cetartiodactyla; Fungivora; Bovidae;		
OC	Bovidae; Bovinae; Tragelaphus.		
OX	NCBI_TaxID=9446;		
KN	[1]		
RP	SEQUENCE.		
PX	MEDLINE: 85279693; PubMed 4026493;		
KA	Rodwald K., Wiesner H., Brautinger G.;		
RT	"Primary structure of the hemoglobin from the Greater Kudu antelope		
RT	(Tragalaphus streptoceros)."		
RL	Biol. Chem. Heppe-Seyler 366:195-402(1985).		
DR	PIR: A02293; HAH KA.		
DR	RSP: P01966; HBA.		
DR	InterPro: IPR004711;		
DR	InterPro: IPR002438;		
DR	PIRfam: PF00042; Globin, 1.		
DR	PRINTS: PR00612; ALPHABEM.		
DR	PROSITE: PS01033; GLOBIN, 1.		
KM	Swiss: 02994; Unisprot; Recombinant; Eukaryotic; Eukaryotic.		
PT	Metab.		
FT	Metab.		
SO	SEQUENCE 140 AA. 1449 MK. A401FA49F43724 CKE4;		
Query Match	100.0%; Score 72; DB 1; Length 140;		
Posterior Similarity	100.0%; Prob 2.8e-06;		
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 PFHFDLSHSAQV 13		
DB	42 PFHFDLSHSAQV 54		
RESULT 2	HBA_TFAST	STANDARD	PRT: 141 AA.
ID	HBA_TFAST		
AC	P11755		
DT	01-OCT-1989 (Rel. 12, Created)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	01-JUN-1994 (Rel. 29, Last annotation update)		
DE	HEMORRHOIDIN ALPHA-1 CHAIN.		
OS	Tadarida brasiliensis (Brazilian free tailed bat);		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eulalia; Chiroptera; Microchiroptera; Molossidae;		
OC	Tadarida.		
OX	NCBI_TaxID=9438;		
KN	[1]		


```

CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: A62265; HAERN.
DR HSP: P01566; LHDA.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHBM.
DR PROSITE: PS01043; GLOBIN_1.
KW Heme oxygen transport; respiratory protein; erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA: 15107 MW: 65420.64 kDa GPC64.

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 2,8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

DB 43 PFBDLSHSAOV 55

RESULT 6
HBA_ANTPA STANDARD: PRT: 141 AA.
AC P14387.
DT 01-JAN-1990 (rel. 13, created)
DI 01-JAN-1990 (rel. 13, last sequence update)
DI 01-OCT-2000 (rel. 40, last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS AUTROZOUS PALLIDUS (Pallid bat).
EC Eukaryota; Metazoa; Chordata; Vertebrata; Eudelostomi;
OC Mammalia; Euphorja; Chiroptera; Microchiroptera; Vespertilionidae;
OC Autrozous.
NCBI_TaxID:9446.
RN [1]
RP SEQUENCE.
RX MEDLINE:86950104; PubMed:3675871;
KA Kleinschmidt T., Koop B.F., Braunitzer G.;
RT "The primary structure of the pallid bat (Autrozous pallidus,
RT Chiroptera) hemoglobin."
KL Biol. Chem. Hoppe-Seyler 368:1197-1202(1987).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: A6792; A2972.
DR HSP: P01922; LFDM.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHBM.
DR PROSITE: PS01043; GLOBIN_1.
KW Heme oxygen transport; respiratory protein; erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA: 15107 MW: 65420.64 kDa GPC64.

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 2,8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

DB 43 PFBDLSHSAOV 55

RESULT 7
HBA_ANTPA STANDARD: PRT: 141 AA.
AC P14387.
DT 01-JAN-1990 (rel. 13, created)
DI 01-JAN-1990 (rel. 13, last sequence update)
DI 01-OCT-2000 (rel. 40, last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS AUTROZOUS PALLIDUS (Pallid bat).
EC Eukaryota; Metazoa; Chordata; Vertebrata; Eudelostomi;
OC Mammalia; Euphorja; Chiroptera; Microchiroptera; Vespertilionidae;
OC Autrozous.
NCBI_TaxID:9446.
RN [1]
RP SEQUENCE.
RX MEDLINE:86950104; PubMed:3675871;
KA Kleinschmidt T., Koop B.F., Braunitzer G.;
RT "The primary structure of the pallid bat (Autrozous pallidus,
RT Chiroptera) hemoglobin."
KL Biol. Chem. Hoppe-Seyler 368:1197-1202(1987).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: A6792; A2972.
DR HSP: P01922; LFDM.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHBM.
DR PROSITE: PS01043; GLOBIN_1.
KW Heme oxygen transport; respiratory protein; erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA: 15107 MW: 65420.64 kDa GPC64.

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 2,8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

DB 43 PFBDLSHSAOV 55

RESULT 8
HBA_ANTPA STANDARD: PRT: 141 AA.
AC P14387.
DT 01-MAR-1989 (rel. 10, created)
DT 01-MAR-1989 (rel. 10, last sequence update)
DI 01-OCT-2000 (rel. 40, last annotation update)
DE HEMOGLOBIN ALPHA-1 AND ALPHA-11 CHAINS.
OS HOMO SAPIENS (Homo sapiens).
EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euphorja; Chordata; Eumetazoa; Eumetazoa; Eumetazoa;
OC Bovidae; Bovinae; Bison.
NCBI_TaxID:9902.
RN [1]
RP SEQUENCE.
RX MEDLINE:86296179; PubMed:3741621;

```

```

HBA_ANTPA STANDARD: PRT: 141 AA.
AC P01927.
DT 21-JUL-1986 (rel. 01, created)
DI 21-JUL-1986 (rel. 01, last sequence update)
DI 01-OCT-2000 (rel. 40, last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS Ateles geoffroyi (black handed spider monkey), and
OS Saguinus oedipus (cotton-top tamarin).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euphorja; Primates; Platyrrhini; Cebellinae; Ateles;
OC NCBI_TaxID:9509, 9490.
RN [1]
RP SEQUENCE.
RX SPECIES-A.geoffroyi;
RX MEDLINE:7509126; PubMed:4215720;
KA Matsuda G., Maita T., Suezawa Y., Setoguchi M., Ota Y., Araya A.,
KA Goodman M., Barnabas J., Prychodko W.;
RT "Studies on the primary structures of alpha and beta polypeptide
RT chains of adult hemoglobin of the spider monkey (Ateles geoffroyi)."
KL Hoppe Seyler's Z. Physik. Chem. 351:1517-1520(1973).
RN [2]
RP SEQUENCE.
RX SPECIES-S.oedipus;
RX MEDLINE:84232683; PubMed:6427302;
KA Maita T., Hayashida M., Matsuda G.;
RT "Primary structures of adult hemoglobins of silver marmoset,
RT Callithrix jacchus, and cotton-headed tamarin, Saguinus oedipus."
KL J. Biochem. 95:855-813(1984).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: A62253; HAEM.
DR PIR: A62253; HAEM.
DR HSP: P01922; LFDM.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHBM.
DR PROSITE: PS01043; GLOBIN_1.
KW Heme oxygen transport; respiratory protein; erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA: 15094 MW: 64925.67 kDa GPC64.

Query Match 100.0%; Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 2,8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

DB 43 PFBDLSHSAOV 55

RESULT 9
HBA_ANTPA STANDARD: PRT: 141 AA.
AC P01927.
DT 01-MAR-1989 (rel. 10, created)
DT 01-MAR-1989 (rel. 10, last sequence update)
DI 01-OCT-2000 (rel. 40, last annotation update)
DE HEMOGLOBIN ALPHA-1 AND ALPHA-11 CHAINS.
OS HOMO SAPIENS (Homo sapiens).
EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euphorja; Chordata; Eumetazoa; Eumetazoa; Eumetazoa;
OC Bovidae; Bovinae; Bison.
NCBI_TaxID:9902.
RN [1]
RP SEQUENCE.
RX MEDLINE:86296179; PubMed:3741621;

```


87 SEQUENCE.
 88 SPECIES: SHEEP: STRAIN: RAMBOUTLET: BREED:
 89 MEDLINE: 660124 PubMed 660045
 90 Wilson J M, Bryant J, Huisman I H, et
 91 "The structure of sheep hemoglobin A."
 92 alpha chain of hemoglobin A."
 93 J Mol Biol. 1966; 24(2):607-652(1966)
 94 151
 95 VARIANT ALLELE:
 96 SEQUENCE: sheep
 97 MEDLINE: 660045 PubMed 660045
 98 Wilson J M, Bryant J, Huisman I H, Vasker P J
 99 "Sheep hemoglobin A: an alpha chain variant with one apparent amino
 100 acid substitution (Glu115 to Asp115)."
 101 Biochem Biophys Acta 1967; 45(1967)
 102 151
 103 1. POLYMERASE CHAIN REACTION DIFFERS FROM THE ALPHA A ALLELE
 104 IN HAVING 75 ASP
 105 1. MULTILANDED: ALLELE: SHEEP: AND OTHER MAMMALIAN SPECIES
 106 PRODUCE IDENTICAL AMINO ACID SEQUENCES FROM THEIR P-NALLET
 107 loci
 108 1. MULTILANDED: THE ALPHA 1 SEQUENCE IS SHOWN
 109 MULTILANDED: BARBARY SHEEP HAS TWO N-TERMINAL ALPHA CHAINS THAT
 110 DIFFER IN THREE POSITIONS: BOTH CHAINS DIFFER FROM THE
 111 GOAT N-TERMINAL CHAIN IN HAVING 18 INSTEAD OF 18 AT POSITION
 112 22 OF 75. THE ALPHA 1 CHAIN IS OTHERWISE APPARENTLY IDENTICAL
 113 WITH THE GOAT CHAIN SHOWN.
 114 MULTILANDED: SHEEP: THE SHEEP ARE EITHER (1) IDENTICAL TO THE 151
 115 GOAT N-TERMINAL CHAIN, OR MORE TWO, AND TWO TIMES MORE THAN THE
 116 GOAT N-TERMINAL CHAIN (151).
 117 ONE MORE THE AND ONE LESS SE.
 118
 119 This Swiss Prot entry is copyright. It is produced through a collaboration
 120 between the Swiss Institute of Bioinformatics and the EMBL outstation
 121 at the European Bioinformatics Institute. There are no restrictions on its
 122 use. By non-profit institutions as long as its content is in no way
 123 modified and this statement is not removed, altered, or in any way
 124 modified to force a certain view on the user. It is not possible to
 125 send an email to EMBL (swissprot@ebi.ac.uk).
 126 EMBL: 151 AA: AAA00001.1
 127 EMBL: 151 AA: AAA00001.1
 128 EMBL: 151 AA: AAA00001.1
 129 EMBL: 151 AA: AAA00001.1
 130 EMBL: 151 AA: AAA00001.1
 131 EMBL: 151 AA: AAA00001.1
 132 EMBL: 151 AA: AAA00001.1
 133 EMBL: 151 AA: AAA00001.1
 134 EMBL: 151 AA: AAA00001.1
 135 EMBL: 151 AA: AAA00001.1
 136 EMBL: 151 AA: AAA00001.1
 137 EMBL: 151 AA: AAA00001.1
 138 EMBL: 151 AA: AAA00001.1
 139 EMBL: 151 AA: AAA00001.1
 140 EMBL: 151 AA: AAA00001.1
 141 EMBL: 151 AA: AAA00001.1
 142 EMBL: 151 AA: AAA00001.1
 143 EMBL: 151 AA: AAA00001.1
 144 EMBL: 151 AA: AAA00001.1
 145 EMBL: 151 AA: AAA00001.1
 146 EMBL: 151 AA: AAA00001.1
 147 EMBL: 151 AA: AAA00001.1
 148 EMBL: 151 AA: AAA00001.1
 149 EMBL: 151 AA: AAA00001.1
 150 EMBL: 151 AA: AAA00001.1

Query Match: Score: 7.21, E-Val: 1.0e-14, Length: 141
 Best Local Similarity: 100.00, Pred. No. Z-Mo: 0.6
 Matches: 13, Conserved Pse: 0, Mismatch: 0, Indels: 0, Gaps: 0
 1. PHEIDISBNAV 13
 14 43 PHEIDISBNAV 55

RESULT 14
 HHA_CERVA STANDARD: PRT: 141 AA.
 11 HHA_CERVA
 AC P01928:
 12 21-JUL-1986 (rev. 01, Created)
 13 21-JUL-1986 (rev. 01, Last sequence update)
 14 01-JUN-1994 (rev. 29, Last annotation update)
 15 01-JUN-1994 (rev. 29, Last annotation update)
 16 HEMOglobIN ALPHAI CHAIN.
 17 Obois apella (brown capoid caputrin).
 18 Eukaryotic Metazoa: Chordata: Vertebrata: Euteleostomi:
 19 Mammalia: Eutheria: Primates: Platyrrhini: Cebidae: Cebus.
 20 NCRL_T010 9516.
 21 111
 22 SEQUENCE:
 23 MEDLINE: 75059125 PubMed 4115717
 24 Matsuda G, Mita T, Watanabe K, Ota Y,
 25 Goodman M, Barndas J, Pychko W,
 26 "The amino acid sequences of the alpha and beta polypeptide chains of
 27 adult hemoglobin A: physical, chemical, and biological properties."
 28 J Biol Chem. 1975; 250(1):151-157(1975).
 29 EMBL: 151 AA: AAA00001.1
 30 EMBL: 151 AA: AAA00001.1
 31 EMBL: 151 AA: AAA00001.1
 32 EMBL: 151 AA: AAA00001.1
 33 EMBL: 151 AA: AAA00001.1
 34 EMBL: 151 AA: AAA00001.1
 35 EMBL: 151 AA: AAA00001.1
 36 EMBL: 151 AA: AAA00001.1
 37 EMBL: 151 AA: AAA00001.1
 38 EMBL: 151 AA: AAA00001.1
 39 EMBL: 151 AA: AAA00001.1
 40 EMBL: 151 AA: AAA00001.1
 41 EMBL: 151 AA: AAA00001.1
 42 EMBL: 151 AA: AAA00001.1
 43 EMBL: 151 AA: AAA00001.1
 44 EMBL: 151 AA: AAA00001.1
 45 EMBL: 151 AA: AAA00001.1
 46 EMBL: 151 AA: AAA00001.1
 47 EMBL: 151 AA: AAA00001.1
 48 EMBL: 151 AA: AAA00001.1
 49 EMBL: 151 AA: AAA00001.1
 50 EMBL: 151 AA: AAA00001.1
 51 EMBL: 151 AA: AAA00001.1
 52 EMBL: 151 AA: AAA00001.1
 53 EMBL: 151 AA: AAA00001.1
 54 EMBL: 151 AA: AAA00001.1
 55 EMBL: 151 AA: AAA00001.1
 56 EMBL: 151 AA: AAA00001.1
 57 EMBL: 151 AA: AAA00001.1
 58 EMBL: 151 AA: AAA00001.1
 59 EMBL: 151 AA: AAA00001.1
 60 EMBL: 151 AA: AAA00001.1
 61 EMBL: 151 AA: AAA00001.1
 62 EMBL: 151 AA: AAA00001.1
 63 EMBL: 151 AA: AAA00001.1
 64 EMBL: 151 AA: AAA00001.1
 65 EMBL: 151 AA: AAA00001.1
 66 EMBL: 151 AA: AAA00001.1
 67 EMBL: 151 AA: AAA00001.1
 68 EMBL: 151 AA: AAA00001.1
 69 EMBL: 151 AA: AAA00001.1
 70 EMBL: 151 AA: AAA00001.1
 71 EMBL: 151 AA: AAA00001.1
 72 EMBL: 151 AA: AAA00001.1
 73 EMBL: 151 AA: AAA00001.1
 74 EMBL: 151 AA: AAA00001.1
 75 EMBL: 151 AA: AAA00001.1
 76 EMBL: 151 AA: AAA00001.1
 77 EMBL: 151 AA: AAA00001.1
 78 EMBL: 151 AA: AAA00001.1
 79 EMBL: 151 AA: AAA00001.1
 80 EMBL: 151 AA: AAA00001.1
 81 EMBL: 151 AA: AAA00001.1
 82 EMBL: 151 AA: AAA00001.1
 83 EMBL: 151 AA: AAA00001.1
 84 EMBL: 151 AA: AAA00001.1
 85 EMBL: 151 AA: AAA00001.1
 86 EMBL: 151 AA: AAA00001.1
 87 EMBL: 151 AA: AAA00001.1
 88 EMBL: 151 AA: AAA00001.1
 89 EMBL: 151 AA: AAA00001.1
 90 EMBL: 151 AA: AAA00001.1
 91 EMBL: 151 AA: AAA00001.1
 92 EMBL: 151 AA: AAA00001.1
 93 EMBL: 151 AA: AAA00001.1
 94 EMBL: 151 AA: AAA00001.1
 95 EMBL: 151 AA: AAA00001.1
 96 EMBL: 151 AA: AAA00001.1
 97 EMBL: 151 AA: AAA00001.1
 98 EMBL: 151 AA: AAA00001.1
 99 EMBL: 151 AA: AAA00001.1
 100 EMBL: 151 AA: AAA00001.1

Query Match 100.0%; Score 72; DB 1; Length 141;
 Best Local Similarity 100.0%; Pred. No. 2.8e+06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FPHFDLSHNSAQN 13
 |||||
 DB 43 FPHFDLSHNSAQN 55

Search completed: June 21, 2001, 10:29:19
 Job time: 248 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - Protein search, using sw model

Run on: June 21, 2001, 10:24:48 : Search time 30.13 seconds
(without alignments)

30.181 Million cell updates/sec

Title: US-08-832-443b-2

Perfect score: 1 CFFPFLSHSAGVC 15

Sequence: HUSIM62
Gap: 10.0, Gapext 0.5

Scoring table: 412676 seqs, 6062398 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Listing first 45 summaries

Database:

A: Geneseq_0601 **
1: /SUS8/qcdata/qcseq/genes/AA1980.DAT *
2: /SUS8/qcdata/qcseq/genes/AA1981.DAT *
3: /SUS8/qcdata/qcseq/genes/AA1982.DAT *
4: /SUS8/qcdata/qcseq/genes/AA1983.DAT *
5: /SUS8/qcdata/qcseq/genes/AA1984.DAT *
6: /SUS8/qcdata/qcseq/genes/AA1985.DAT *
7: /SUS8/qcdata/qcseq/genes/AA1986.DAT *
8: /SUS8/qcdata/qcseq/genes/AA1987.DAT *
9: /SUS8/qcdata/qcseq/genes/AA1988.DAT *
10: /SUS8/qcdata/qcseq/genes/AA1989.DAT *
11: /SUS8/qcdata/qcseq/genes/AA1990.DAT *
12: /SUS8/qcdata/qcseq/genes/AA1991.DAT *
13: /SUS8/qcdata/qcseq/genes/AA1992.DAT *
14: /SUS8/qcdata/qcseq/genes/AA1993.DAT *
15: /SUS8/qcdata/qcseq/genes/AA1994.DAT *
16: /SUS8/qcdata/qcseq/genes/AA1995.DAT *
17: /SUS8/qcdata/qcseq/genes/AA1996.DAT *
18: /SUS8/qcdata/qcseq/genes/AA1997.DAT *
19: /SUS8/qcdata/qcseq/genes/AA1998.DAT *
20: /SUS8/qcdata/qcseq/genes/AA1999.DAT *
21: /SUS8/qcdata/qcseq/genes/AA2000.DAT *
22: /SUS8/qcdata/qcseq/genes/AA2001.DAT *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	15	AAW0316	Cyclic peptide 43-
2	79.5	88.3	16	AAK4516	Peptide fragment 9
3	72	80.0	13	AAW0330	Reptide 43-55 with
4	72	80.0	16	AAV1983	Rheumatoid arthritis
5	72	80.0	14	AAW6276	Circulating peptide
6	72	80.0	14	AAK2960	Mutant human alpha
7	72	80.0	14	AAK2960	Mutant human alpha
8	72	80.0	14	AAK1053	Titusville alpha 9
9	72	80.0	14	AAK4298	Human alpha haemoglobin
10	72	80.0	14	AAK4298	Natural alpha globulin
11	72	80.0	14	AAK5131	Human alpha haemoglobin

12	72	80.0	14	AAK44516	Adult human alpha
13	72	80.0	14	AAK44517	Adult human alpha
14	72	80.0	14	AAK44518	Adult human alpha
15	72	80.0	14	AAK44519	Adult human alpha
16	72	80.0	14	AAK44520	Adult human alpha
17	72	80.0	14	AAK47359	Human alpha haemoglobin
18	72	80.0	14	AAW03312	The alpha chain of
19	72	80.0	14	AAW0734	Haemoglobin alpha
20	72	80.0	14	AAW0734	Haemoglobin alpha
21	72	80.0	14	AAW0734	Native human haemoglobin
22	72	80.0	14	AAW0724	Haemoglobin alpha
23	72	80.0	14	AAW0725	Haemoglobin alpha
24	72	80.0	14	AAW0726	Haemoglobin alpha
25	72	80.0	14	AAW0716	Recombinant wild type
26	72	80.0	14	AAW10470	Human alpha-haemoglobin
27	72	80.0	14	AAW11868	Human alpha-haemoglobin
28	72	80.0	14	AAW7389	Human haemoglobin
29	72	80.0	14	AAW7386	Human haemoglobin
30	72	80.0	14	AAW0505	Alpha chain of human
31	72	80.0	14	AAW03376	Human alpha-haemoglobin
32	72	80.0	14	AAW04541	Human alpha-haemoglobin
33	72	80.0	14	AAW07794	Human alpha-haemoglobin
34	72	80.0	14	AAW79379	Human mutant alpha
35	72	80.0	14	AAW65774	Human haemoglobin
36	72	80.0	14	AAW06058	Human haemoglobin
37	72	80.0	14	AAK19457	(Prox-Val) alpha (6
38	72	80.0	14	AAW54292	Protein sequence of
39	72	80.0	14	AAW08114	(Prox-Val) alpha (6
40	72	80.0	14	AAW64283	DNA sequence of the
41	72	80.0	14	AAW54284	DNA sequence of the
42	72	80.0	14	AAW54219	Human adult alpha
43	72	80.0	14	AAW54219	Human adult alpha
44	72	80.0	14	AAW54219	Human adult alpha
45	72	80.0	14	AAW70979	Low oxygen affinity

ALIGNMENTS

RESULT 1	
AAW0316	
ID	AAW0316 standard; peptide: 15 AA.
XX	
AC	AAW0316:
XX	
XX	20-APR-1998 (first entry)
XX	
XX	Cyclic peptide 43-55 with stem cell inhibitory/stimulatory activity.
XX	
XX	INPRED stem cell inhibitory stimulation: proliferation; myoglobin
XX	alpha chain; myoglobin; adult haemoglobin; alpha 9
XX	differentiated production; stem cell hypoproliferation; apoptotic anemias
XX	stem cell exhaustion; cyclic peptide.
XX	
XX	Synthetic.
XX	Homo sapiens.
XX	
XX	Key
XX	Disulfide bond 1..15
XX	
XX	MO9746922-A1.
XX	
XX	09-0CT-1997.
XX	
XX	03-APR-1997: 97MO-US05601.
XX	
XX	03-APR-1997: 97MO-US05601.
XX	
XX	(FROM) PRO NEBROX INC.
XX	
XX	TSYLOVA I, Wolpe SD:
XX	
XX	MD1: 1997-503044/46.

DR WP1: 1997-50044/46.
 XX
 PS Haemoglobin alpha chain derivatives - used to inhibit or stimulate
 PT stem cell proliferation, for treatment of cancer, immune depression
 PT etc
 XX
 PS Claim 46: Page 116: 162pp; English.
 CC The present peptide is an INP601 peptide that has stimulatory and
 CC inhibitory activity on stem cell proliferation. The 3-dimensional
 CC structure of myoglobin was superimposed onto the 3-dimensional
 CC structure of the alpha chain, present in human haemoglobin, and the
 CC present peptide identified as a region which is structurally different
 CC from myoglobin in 3-dimensional space. The peptide is used to
 CC inhibit stem cell proliferation, particularly during radiotherapy or
 CC chemotherapy of cancer, but more generally wherever a mammal is exposed
 CC to an agent, e.g. an antiviral, that damages or destroys stem cells. It
 CC can also be used for differential protection of normal stem cells, but
 CC not cancer cells, from chemotherapy or radiation, particularly after
 CC normal stem cells have been induced to proliferate by therapeutic
 CC treatment. The peptide can also be used for control of stem cell
 CC hyperproliferation (e.g. aplastic anaemia), for treating or preventing
 CC stem cell exhaustion (e.g. where caused by acquired immune deficiency
 CC syndrome), and to treat or prevent immunodeficiency. INP601 can also be
 CC used for treatment of a mammal, adult and related compounds can
 CC reversibly inhibit or stimulate stem cells depending on the dose, so
 CC allow precise control over cycling of these cells.
 XX
 S0 Sequence 1: AA:
 Query Match: 80.0%; Score 72; EB 18; Length 13;
 Best Local Similarity: 100.0%; Pos. No. 2,28-04;
 Matches: 13; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 QY 2 FHHHDLSSNAV 14
 DB 1 FHHHDLSSNAV 14
 RESULT 4
 AAV41983
 ID AAV41983 standard; Peptide: 16 AA.
 AC AAV41983;
 XX
 DT 09-DEC-1999 (first entry)
 XX
 DE Rheumatoid arthritis diagnostic protein isoform peptide #14
 XX
 KW human; rheumatoid arthritis; RA; diagnosis; RAI; RAIPE; detection;
 KW rheumatoid arthritis diagnostic feature; RAIPE; synovial fluid;
 KW rheumatoid arthritis diagnostic protein isoform; sequencing;
 KW expression reference protein isoform; prognosis.
 XX
 OS Homo sapiens.
 XX
 PN W09947925-A2.
 XX
 PD 21-SEP-1999.
 XX
 PF 15-MAR-1999; GWO-GB00763.
 XX
 PR 13-MAR-1998; GSGR-0005477.
 XX
 PA (OXFO-) XPORD GLYC/SCIENCES UK LTD.
 XX
 PT Parekh RB, Patel TP, Townsend RB;
 XX
 DR WP1: 1999-571871/48
 XX
 PT diagnosis of human rheumatoid arthritis by two-dimensional
 PT electrophoresis -

XX
 PS Disclosure: Page 26: 157pp; English.
 XX
 CC A method has been developed for the diagnosis of human rheumatoid
 CC arthritis (RA) using two-dimensional electrophoresis to generate a
 CC two dimensional array of features. The method can be used for screening
 CC diagnosis and prognosis of RA in a subject or for monitoring the effect
 CC of an anti-RA drug or therapy administered to a subject. The method
 CC comprises: (a) analysing a sample of serum or plasma and optionally
 CC synovial fluid by two dimensional electrophoresis, to generate a two-
 CC dimensional array of features; (b) identifying at least one chosen
 CC feature whose relative abundance correlates with the presence or absence
 CC of RA; and (c) comparing the abundance of each chosen feature in the
 CC sample with the abundance of that chosen feature in serum or plasma from
 CC one or more persons without RA, where the relative abundance of the
 CC chosen feature or features in the sample indicates the presence or
 CC absence of RA in the subject. The method can also be used in clinical
 CC studies for testing drugs for therapy of RA, for purification of RA-
 CC diagnostic protein isoforms (RPIs), and for production of antibodies to
 CC RPIs. The RA diagnostic feature (RAUF) proteins can be used to identify
 CC compounds that promote or inhibit their activity, which are then used as
 CC RA drugs. Nucleic acid encoding RAIPEs can be used in gene therapy.
 CC Proteolysis: AAV41944 to AAV42103 represent RPI peptides; AAV42101 to
 CC AAV42103 represent expression reference protein isoform peptides and
 CC AAV42008 to AAV42092 represent degenerate protein RPIs, which are all
 CC used in the exemplification of the present invention.
 XX
 S0 Sequence 16 AA:
 Query Match: 80.0%; Score 72; DB 20; Length 16;
 Best Local Similarity: 100.0%; Pos. No. 2,78-06;
 Matches: 13; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 QY 2 FHHHDLSSNAV 14
 DB 3 FHHHDLSSNAV 14
 RESULT 5
 AAW66275
 ID AAW66275 standard; peptide: 19 AA.
 AC AAW66275;
 XX
 DT 23-NOV-1998 (first entry)
 XX
 DE circularly permuted alpha haemoglobin-like peptide fragment.
 XX
 KW haemoglobin; circularly permuted; N terminus; C terminus; multi-mer;
 KW permuted haemoglobin; globin; pHb; red blood cells; blood transfusion;
 KW sickle cells; stroke.
 XX
 OS Synthetic.
 XX
 PN W09848211-A2.
 XX
 PD 03-SEP-1998.
 XX
 PF 27-FEB-1998; 98WO-050836.
 XX
 PR 28-FEB-1997; 97US-0039657.
 XX
 FA (SOMA-) SOMALOGIN INC.
 XX
 PT Best EA, Collins PO;
 XX
 DR WP1: 1998-481141/41.
 XX
 PT New haemoglobin formulation with non-naturally occurring termini
 PT useful for the treatment of any condition where therapy involves red
 PT blood cells or oxygen delivery
 XX

101 the transfusion-related to a haemoglobin form (Hb) containing at
 102 least one globin with non naturally occurring N and C termini. Also
 103 claimed is a genetic acid encoding the above globin. Hb can be used as a
 104 substrate for red blood cells phenotypic and blood cells or oxygen
 105 delivery is required, particularly in the treatment of haemophobias,
 106 trauma, and various conditions requiring blood transfusion or retention of the
 107 body's own blood and anemia, to deliver oxygen to tissue sites that
 108 need blood cells cannot penetrate, including sites distant from an
 109 artery or vein, stroke, and cerebral malformation, and to treat tissue
 110 ischaemia, including stroke, myocardial infarction, and to treat tissue
 111 myocardial infarction and haemophilia, or other unstable condition, and
 112 in addition, the polymers have increased stability and expression during in
 113 vivo synthesis and a structure in vivo equivalent half-life compared to
 114 the parent haemoglobin. The present sequence is shown

115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

1001 mutation was then added to the sequence. The sequence was mutated by
 1002 polymerase chain reaction. The mutation introduced produced a
 1003 exchange characteristic sufficiently close to natural haemoglobin (Hb)
 1004 to enable its use in a cell free blood substitute. The mutated
 1005 gene encoding this protein can be used in a construct designed for
 1006 expression of pure, modified human Hb's in the erythrocytes of non-
 1007 human transgenic animals. The Hb's are non immunogenic so can be used
 1008 without the need for blood typing. Also, the Transgene Hb will be
 1009

1010
 1011
 1012
 1013
 1014
 1015
 1016
 1017
 1018
 1019
 1020
 1021
 1022
 1023
 1024
 1025
 1026
 1027
 1028
 1029
 1030
 1031
 1032
 1033
 1034
 1035
 1036
 1037
 1038
 1039
 1040
 1041
 1042
 1043
 1044
 1045
 1046
 1047
 1048
 1049
 1050
 1051
 1052
 1053
 1054
 1055
 1056
 1057
 1058
 1059
 1060
 1061
 1062
 1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072
 1073
 1074
 1075
 1076
 1077
 1078
 1079
 1080
 1081
 1082
 1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133
 1134
 1135
 1136
 1137
 1138
 1139
 1140
 1141
 1142
 1143
 1144
 1145
 1146
 1147
 1148
 1149
 1150
 1151
 1152
 1153
 1154
 1155
 1156
 1157
 1158
 1159
 1160
 1161
 1162
 1163
 1164
 1165
 1166
 1167
 1168
 1169
 1170
 1171
 1172
 1173
 1174
 1175
 1176
 1177
 1178
 1179
 1180
 1181
 1182
 1183
 1184
 1185
 1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457
 1458
 1459
 1460
 1461
 1462
 1463
 1464
 1465
 1466
 1467
 1468
 1469
 1470
 1471
 1472
 1473
 1474
 1475
 1476
 1477
 1478
 1479
 1480
 1481
 1482
 1483
 1484
 1485
 1486
 1487
 1488
 1489
 1490
 1491
 1492
 1493
 1494
 1495
 1496
 1497
 1498
 1499
 1500

1501
 1502
 1503
 1504
 1505
 1506
 1507

PN W09325071-A.
 XX
 PD 23-DEC-1994.
 XX
 XX 11-JUN-1994: 94WO-0805629.
 XX
 PF 12-JUN-1992: 92US-0897648.
 PR 08-DEC-1992: 92US-0987890.
 PR 15-MAR-1993: 94US-0030897.
 XX
 PA (DNXP-) DNK CORP.
 XX
 P1 Holtzman SM, Keller R, Kumar R, Logan JS, O'Connell JK,
 P1 Parsons CT, Pinder SH, Pinkert CA, Sharma A, Swanson ME,
 P1 White SP;
 DR WPI: 1994-007075/01.
 XX
 PT Transgenic pigs which produce human haemoglobin - cont'd. DNA encoding
 PT human alpha globin and huma beta globin linked to promoters
 PS
 PS Claim 14: : 117pp; English.
 XX
 CC The sequences given in AAR44516-20 represent mutant versions of the
 CC adult human alpha globin sequence. These sequences were not given in
 CC the specification but were generated using the known alpha globin
 CC sequence. These sequences were used in the production of constructs which
 CC were used to produce the transgenic pigs of the invention. These
 CC constructs contain the pig beta globin promoter sequence operably linked
 CC to the human alpha or beta globin genes. Pigs containing these
 CC constructs express human haemoglobin (Hb) in their erythrocytes and are
 CC healthy, suffering no deleterious side effects as a result of
 CC heterologous Hb production. They can be used as an efficient and
 CC economical source of human Hb that can be used for transfusion and other
 CC medical applications. Also, a human/pig hybrid Hb can be produced which
 CC exhibits a p50 that is higher than that of native human or pig Hb.
 CC
 SO Sequence 141 AA:
 QY
 Query Match 80.0%; Score 72; DB 15; Length 141;
 Best Local Similarity 100.0%; Pred. No. 30-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 DB 43 fphdlsbasy 55
 QY 2 FPHDLSBSACV 14
 ID AAR44517
 AC AAR44517 standard: Protein; 141 AA.
 XX
 AC AAR44517;
 XX
 DE 23-JUN-1994 (first entry)
 XX
 DE Adult human alpha-globin F36Y.
 XX
 KW Adult; pig; beta; globin gene; regulatory region; promoter;
 KW transgenic pig; human; alpha; haemoglobin; Hb; erythroytic;
 KW transgenic; human/pig; hybrid Hb.
 XX
 OS Homo sapiens.
 XX
 FT Key local quality factors
 FT Misc-difference 46 /Label- F36Y
 XX
 PN W09325071-A.
 XX
 PD 23-DEC-1994.
 XX

PF 11-JUN-1994: 94WO-0805629.
 XX
 PR 12-JUN-1992: 92US-0897648.
 PR 08-DEC-1992: 92US-0987890.
 PR 15-MAR-1993: 94US-0030897.
 XX
 PA (DNXP-) DNK CORP.
 XX
 P1 Holtzman SM, Keller R, Kumar R, Logan JS, O'Connell JK,
 P1 Parsons CT, Pinder SH, Pinkert CA, Sharma A, Swanson ME,
 P1 White SP;
 DR WPI: 1994-007075/01.
 XX
 PT Transgenic pigs which produce human haemoglobin - cont'd. DNA encoding
 PT human alpha globin and huma beta globin linked to promoters
 PS
 PS Claim 14: : 117pp; English.
 XX
 CC The sequences given in AAR44516-20 represent mutant versions of the
 CC adult human alpha globin sequence. These sequences were not given in
 CC the specification but were generated using the known alpha globin
 CC sequence. These sequences were used in the production of constructs which
 CC were used to produce the transgenic pigs of the invention. These
 CC constructs contain the pig beta globin promoter sequence operably linked
 CC to the human alpha or beta globin genes. Pigs containing these and are
 CC healthy, suffering no deleterious side effects as a result of
 CC heterologous Hb production. They can be used as an efficient and
 CC economical source of human Hb that can be used for transfusion and other
 CC medical applications. Also, a human/pig hybrid Hb can be produced which
 CC exhibits a p50 that is higher than that of native human or pig Hb.
 CC
 SO Sequence 141 AA:
 QY
 Query Match 80.0%; Score 72; DB 15; Length 141;
 Best Local Similarity 100.0%; Pred. No. 30-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 DB 43 fphdlsbasy 55
 QY 2 FPHDLSBSACV 14
 ID AAR44518
 AC AAR44518 standard: Protein; 141 AA.
 XX
 AC AAR44518;
 XX
 DE 23-JUN-1994 (first entry)
 XX
 DE Adult human alpha-globin L106F.
 XX
 KW Adult; pig; beta; globin gene; regulatory region; promoter;
 KW transgenic pig; human; alpha; haemoglobin; Hb; erythroytic;
 KW transgenic; human/pig; hybrid Hb.
 XX
 OS Homo sapiens.
 XX
 FT Key local quality factors
 FT Misc-difference 106 /Label- L106F
 XX
 PN W09325071-A.
 XX
 PD 23-DEC-1994.
 XX
 PF 11-JUN-1994: 94WO-0805629.
 XX
 PR 12-JUN-1992: 92US-0897648.
 PR 08-DEC-1992: 92US-0987890.
 XX

[illegible][illegible]

1	RESIDUE	1			
2	CONFID				
3	1	CONFID			
4	2	CONFID			
5	3	CONFID			
6	4	CONFID			
7	5	CONFID			
8	6	CONFID			
9	7	CONFID			
10	8	CONFID			
11	9	CONFID			
12	10	CONFID			
13	11	CONFID			
14	12	CONFID			
15	13	CONFID			
16	14	CONFID			
17	15	CONFID			
18	16	CONFID			
19	17	CONFID			
20	18	CONFID			
21	19	CONFID			
22	20	CONFID			
23	21	CONFID			
24	22	CONFID			
25	23	CONFID			
26	24	CONFID			
27	25	CONFID			
28	26	CONFID			
29	27	CONFID			
30	28	CONFID			
31	29	CONFID			
32	30	CONFID			
33	31	CONFID			
34	32	CONFID			
35	33	CONFID			
36	34	CONFID			
37	35	CONFID			
38	36	CONFID			
39	37	CONFID			
40	38	CONFID			
41	39	CONFID			
42	40	CONFID			
43	41	CONFID			
44	42	CONFID			
45	43	CONFID			
46	44	CONFID			
47	45	CONFID			
48	46	CONFID			
49	47	CONFID			
50	48	CONFID			
51	49	CONFID			
52	50	CONFID			
53	51	CONFID			
54	52	CONFID			
55	53	CONFID			
56	54	CONFID			
57	55	CONFID			
58	56	CONFID			
59	57	CONFID			
60	58	CONFID			
61	59	CONFID			
62	60	CONFID			
63	61	CONFID			
64	62	CONFID			
65	63	CONFID			
66	64	CONFID			
67	65	CONFID			
68	66	CONFID			
69	67	CONFID			
70	68	CONFID			
71	69	CONFID			
72	70	CONFID			
73	71	CONFID			
74	72	CONFID			
75	73	CONFID			
76	74	CONFID			
77	75	CONFID			
78	76	CONFID			
79	77	CONFID			
80	78	CONFID			
81	79	CONFID			
82	80	CONFID			
83	81	CONFID			
84	82	CONFID			
85	83	CONFID			
86	84	CONFID			
87	85	CONFID			
88	86	CONFID			
89	87	CONFID			
90	88	CONFID			
91	89	CONFID			
92	90	CONFID			
93	91	CONFID			
94	92	CONFID			
95	93	CONFID			
96	94	CONFID			
97	95	CONFID			
98	96	CONFID			
99	97	CONFID			
100	98	CONFID			
101	99	CONFID			
102	100	CONFID			
103	101	CONFID			
104	102	CONFID			
105	103	CONFID			
1					

[illegible]

```

007 1 FFFFD.SHOAV 14
      | | | | | | | |
146 44 FFFFD.SHOAV 59

```

[illegible]

01	0001	100.00%	Score 7.2	DB 0.2	Month 14.2	0.2
02	0002	100.00%	Prod. No. 4	90-06		
03	0003	100.00%	01	MSBIRTHS	02	INDCLS
04	0004	100.00%	02	MSBIRTHS	03	INDCLS
05	0005	100.00%	03	MSBIRTHS	04	INDCLS
06	0006	100.00%	04	MSBIRTHS	05	INDCLS
07	0007	100.00%	05	MSBIRTHS	06	INDCLS
08	0008	100.00%	06	MSBIRTHS	07	INDCLS
09	0009	100.00%	07	MSBIRTHS	08	INDCLS
10	0010	100.00%	08	MSBIRTHS	09	INDCLS
11	0011	100.00%	09	MSBIRTHS	10	INDCLS
12	0012	100.00%	10	MSBIRTHS	11	INDCLS
13	0013	100.00%	11	MSBIRTHS	12	INDCLS
14	0014	100.00%	12	MSBIRTHS	13	INDCLS
15	0015	100.00%	13	MSBIRTHS	14	INDCLS
16	0016	100.00%	14	MSBIRTHS	15	INDCLS
17	0017	100.00%	15	MSBIRTHS	16	INDCLS
18	0018	100.00%	16	MSBIRTHS	17	INDCLS
19	0019	100.00%	17	MSBIRTHS	18	INDCLS
20	0020	100.00%	18	MSBIRTHS	19	INDCLS
21	0021	100.00%	19	MSBIRTHS	20	INDCLS
22	0022	100.00%	20	MSBIRTHS	21	INDCLS
23	0023	100.00%	21	MSBIRTHS	22	INDCLS
24	0024	100.00%	22	MSBIRTHS	23	INDCLS
25	0025	100.00%	23	MSBIRTHS	24	INDCLS
26	0026	100.00%	24	MSBIRTHS	25	INDCLS
27	0027	100.00%	25	MSBIRTHS	26	INDCLS
28	0028	100.00%	26	MSBIRTHS	27	INDCLS
29	0029	100.00%	27	MSBIRTHS	28	INDCLS
30	0030	100.00%	28	MSBIRTHS	29	INDCLS
31	0031	100.00%	29	MSBIRTHS	30	INDCLS
32	0032	100.00%	30	MSBIRTHS	31	INDCLS
33	0033	100.00%	31	MSBIRTHS	32	INDCLS
34	0034	100.00%	32	MSBIRTHS	33	INDCLS
35	0035	100.00%	33	MSBIRTHS	34	INDCLS
36	0036	100.00%	34	MSBIRTHS	35	INDCLS
37	0037	100.00%	35	MSBIRTHS	36	INDCLS
38	0038	100.00%	36	MSBIRTHS	37	INDCLS
39	0039	100.00%	37	MSBIRTHS	38	INDCLS
40	0040	100.00%	38	MSBIRTHS	39	INDCLS
41	0041	100.00%	39	MSBIRTHS	40	INDCLS
42	0042	100.00%	40	MSBIRTHS	41	INDCLS
43	0043	100.00%	41	MSBIRTHS	42	INDCLS
44	0044	100.00%	42	MSBIRTHS	43	INDCLS
45	0045	100.00%	43	MSBIRTHS	44	INDCLS
46	0046	100.00%	44	MSBIRTHS	45	INDCLS
47	0047	100.00%	45	MSBIRTHS	46	INDCLS
48	0048	100.00%	46	MSBIRTHS	47	INDCLS
49	0049	100.00%	47	MSBIRTHS	48	INDCLS
50	0050	100.00%	48	MSBIRTHS	49	INDCLS
51	0051	100.00%	49	MSBIRTHS	50	INDCLS
52	0052	100.00%	50	MSBIRTHS	51	INDCLS
53	0053	100.00%	51	MSBIRTHS	52	INDCLS
54	0054	100.00%	52	MSBIRTHS	53	INDCLS
55	0055	100.00%	53	MSBIRTHS	54	INDCLS
56	0056	100.00%	54	MSBIRTHS	55	INDCLS
57	0057	100.00%	55	MSBIRTHS	56	INDCLS
58	0058	100.00%	56	MSBIRTHS	57	INDCLS
59	0059	100.00%	57	MSBIRTHS	58	INDCLS
60	0060	100.00%	58	MSBIRTHS	59	INDCLS
61	0061	100.00%	59	MSBIRTHS	60	INDCLS
62	0062	100.00%	60	MSBIRTHS	61	INDCLS
63	0063	100.00%	61	MSBIRTHS	62	INDCLS
64	0064	100.00%	62	MSBIRTHS	63	INDCLS
65	0065	100.00%	63	MSBIRTHS	64	

RX MEDLINE: 9865770; PubMed: 9847419;
 RA Oakenfull E.A., Clegg J.B.;
 RT "Phylogenetic relationships within the genus Equus and the evolution
 of alpha and theta globin genes.";
 RL J. Mol. Evol. 47:772-783(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Oakenfull E.A., Clegg J.B.;
 RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO GLOBIN FAMILY.
 CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.
 DR EMBL: AF129139; AAB93466.1;
 DR EMBL: AF129138; AAB93466.1; JOINED.
 DR HSSP: P01958; 11BF.
 DR InterPro: IPR000971;
 DR InterPro: IPR002338;
 DR Pfam: PF00042; GLOBIN.1.
 DR PRINTS: PR00612; ALPHAEQM.
 DR PROSITE: PS01033; GLOBIN.1.
 KM Erythrocyte, Hemoglobin, Oxygen transport, Respiratory protein.
 SQ SEQUENCE 142 AA; 15211 MW; 82A452F095949854 CRC64;

Query Match 100.0%; Score 72; DP 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 6; 98-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDSHGSAAV 13
 DB 44 FPHFDSHGSAAV 56
 RESULT 13
 QYTVAA
 ID QYTVAA PRELIMINARY; PRT; 142 AA.
 AC QYTVAA;
 DT 01-MAY-2000 (TRENKREL, 13; Created)
 DT 01-MAY-2000 (TRENKREL, 13; Last sequence update)
 DT 01-MAR-2001 (TRENKREL, 16; Last annotation update)
 DE ALPHA 1 HEMOGLOBIN.
 OS Equus hemionus (Kulun).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eultheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oakenfull E.A., Clegg J.B.;
 RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO GLOBIN FAMILY.
 CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.
 DR EMBL: AF129137; AAB93465.1;
 DR EMBL: AF129136; AAB93465.1; JOINED.
 DR HSSP: P01958; 11BF.
 DR InterPro: IPR000971;
 DR InterPro: IPR002338;
 DR Pfam: PF00042; globin.1
 DR PRINTS: PR00612; ALPHAEQM
 DR PROSITE: PS01033; GLOBIN.1
 KM Erythrocyte, Hemoglobin, Oxygen transport, Respiratory protein.
 SQ SEQUENCE 142 AA; 15206 MW; 82B878C14C1E1B2 CRC64;

Query Match 100.0%; Score 72; DP 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 6; 98-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDSHGSAAV 13
 DB 44 FPHFDSHGSAAV 56

RESULT 14
 QYTVAA3
 ID QYTVAA3 PRELIMINARY; PRT; 142 AA.
 AC QYTVAA3;
 DT 01-MAY-2000 (TRENKREL, 13; Created)
 DT 01-MAY-2000 (TRENKREL, 13; Last sequence update)
 DT 01-MAR-2001 (TRENKREL, 16; Last annotation update)
 DE ALPHA 2 HEMOGLOBIN.
 OS Equus burchelli (Plains zebra) (Equus quagga).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eultheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9790;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oakenfull E.A., Clegg J.B.;
 RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO GLOBIN FAMILY.
 CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.
 DR EMBL: AF129134; AAB93462.1;
 DR EMBL: AF129133; AAB93462.1; JOINED.
 DR HSSP: P01958; 11BF.
 DR InterPro: IPR000971;
 DR InterPro: IPR002338;
 DR Pfam: PF00042; globin.1.
 DR PRINTS: PR00612; ALPHAEQM.
 DR PROSITE: PS01033; GLOBIN.1.
 KM Erythrocyte, Hemoglobin, Oxygen transport, Respiratory protein.
 SQ SEQUENCE 142 AA; 15190 MW; 8E783780801D85 CRC64;

Query Match 100.0%; Score 72; DP 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 6; 98-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPHFDSHGSAAV 13
 DB 44 FPHFDSHGSAAV 56
 RESULT 15
 QYTVS1
 ID QYTVS1 PRELIMINARY; PRT; 142 AA.
 AC QYTVS1;
 DT 01-MAY-2000 (TRENKREL, 13; Created)
 DT 01-MAY-2000 (TRENKREL, 13; Last sequence update)
 DT 01-MAR-2001 (TRENKREL, 16; Last annotation update)
 DE HEMOGLOBIN ALPHA CHAIN.
 GN ALPHA GLOBIN.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eultheria; Perissodactyla; Ruminantia; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID 9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kullio R., Pietraszewska E., Vincenzini D., Campanile G., Di Lucia A.;
 RT "Nucleotide sequence of Fetal human gamma (HbA2) and alpha
 1-globin genes.";
 SQ Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO GLOBIN FAMILY.


```

1  STATE: VA
2  ZIP: 04111
3  COMPUTER READABLE FORM
4  MEDIUM TYPE: floppy disk
5  MEDIUM TYPE: 100 pc compat disk
6  OPERATING SYSTEM: pc dos/ms dos
7  SOFTWARE: word 5.0a
8  CURRENT APPLICATION DATA
9  APPLICATION NUMBER: 05/07/94/443b-2
10 FILING DATE: 11 JUN 1994
11 CLASSIFICATION: 40
12 PRIOR APPLICATION DATA
13 APPLICATION NUMBER: 05/07/94/443b-2
14 FILING DATE: 22 OCT 1990
15 PRIOR APPLICATION DATA
16 FILING DATE: 22 OCT 1990
17 NAME: 443b-2
18 REGISTRATION NUMBER: 44200
19 TELEPHONE: 403 444 4014
20 TELEFAX: 403 444 4014
21 INFORMATION FOR SEQ ID NO: 1:
22 SEQUENCE CHARACTERISTICS
23 LENGTH: 141
24 TYPE: amino acid
25 MODIFICATION: protein
26 07/94/443b-2
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 641,617
FILING DATE: 16-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 310,891
FILING DATE: 17-FEB-1989
PRIOR APP. DATA:
APPLICATION NUMBER: US 160,766
FILING DATE: 26-FEB-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 160,771
FILING DATE: 28-FEB-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 447,637
FILING DATE: 05-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 463,138
FILING DATE: 08-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 219,279
FILING DATE: 15-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: EUS-20121 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-433-4150
TELEFAX: 415-433-8716
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-184-237-8

Query Match: 80.0% Score 72, PB 1, Length 141;
Host Local Similarity: 100.0%, Pred. No. 5,36-05;

Matches: 13, Conservative: 0, Mismatches: 0, Indels: 0, Gaps: 0;

QY 2 FPHDLSHSAV 14
ID 43 FPHDLSHSAV 55
RESULT 10
US-08-240-712-17
Sequence: 17, Application: US/08-44712
Patent No. 5599907
GENERAL INFORMATION:
APPLICANT: ANDERSON, DAVID C.
APPLICANT: MATHERS, ANTONY JAMES
APPLICANT: STEHLER, GARY L.
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Broadway and Newmark
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09-240-712
FILING DATE: 09-MAY-1994
CLASSIFICATION: 5:40

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09752
FILING DATE: 13-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: AMELPSON-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-747-9528
TELEX: 248643
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-240-712-17

Query Match: 80.0% Score 72, PB 1, Length 141;
Host Local Similarity: 100.0%, Pred. No. 5,36-05;

Matches: 13, Conservative: 0, Mismatches: 0, Indels: 0, Gaps: 0;

QY 2 FPHDLSHSAV 14
ID 43 FPHDLSHSAV 55

RESULT 11
US-08-240-712-26

Sequence: 26, Application: US/08-240712
Patent No. 5599907

GENERAL INFORMATION:

APPLICANT: ANDERSON, DAVID C.

APPLICANT: MATHERS, ANTONY JAMES

APPLICANT: STEHLER, GARY L.

TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Broadway and Newmark

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08-240-712

FILING DATE: 09-MAY-1994

CLASSIFICATION: 5:40

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/09752

FILING DATE: 13-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: COOPER, IVER P.

REGISTRATION NUMBER: 28,005

REFERENCE/DOCKET NUMBER: AMELPSON-6

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-747-9528

TELEX: 248643

INFORMATION FOR SEQ. ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 141 amino acids

TYPE: amino acid

STRANDEDNESS: single



GenCode version 4.5
Copyright (c) 1994 - 2000 Compugen Ltd.

OM protein protein search, using sw model

Run on: June 21, 2001, 10:29:19 : Search time 11.96 seconds

(without alignments)
42.663 Million cell updates/sec

File: US-08-832-443b-2

Perfect score: 50

Sequence: 1 FPHFDLSHSAQVC 15

Scoring table: BLSDM62

Gapop 10 0, Gapov 0 5

Searched: 9445 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 94435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39:*

Result No. Score Match Length DB ID

1 72 80.0 141 1 HBA_TRAFT

2 72 80.0 141 1 HBA_TRAFT

3 72 80.0 141 1 HBA_TRAFT

4 72 80.0 141 1 HBA_TRAFT

5 72 80.0 141 1 HBA_TRAFT

6 72 80.0 141 1 HBA_TRAFT

7 72 80.0 141 1 HBA_TRAFT

8 72 80.0 141 1 HBA_TRAFT

9 72 80.0 141 1 HBA_TRAFT

10 72 80.0 141 1 HBA_TRAFT

11 72 80.0 141 1 HBA_TRAFT

12 72 80.0 141 1 HBA_TRAFT

13 72 80.0 141 1 HBA_TRAFT

14 72 80.0 141 1 HBA_TRAFT

15 72 80.0 141 1 HBA_TRAFT

16 72 80.0 141 1 HBA_TRAFT

17 72 80.0 141 1 HBA_TRAFT

18 72 80.0 141 1 HBA_TRAFT

19 72 80.0 141 1 HBA_TRAFT

20 72 80.0 141 1 HBA_TRAFT

Description

P04237 Tagelaphus

P11757 Tagelaphus

P13300 Tagelaphus

P01968 Tagelaphus

P01971 Tagelaphus

P14381 Tagelaphus

P01927 Tagelaphus

P09423 Tagelaphus

P01969 Tagelaphus

P01966 Tagelaphus

P18973 Tagelaphus

P01970 Tagelaphus

P01928 Tagelaphus

P01421 Tagelaphus

P01926 Tagelaphus

P01963 Tagelaphus

P01930 Tagelaphus

P18973 Tagelaphus

P01970 Tagelaphus

ALIGNMENTS

34 72 80.0 141 1 HBA_MACAS
35 72 80.0 141 1 HBA_MACAS
36 72 80.0 141 1 HBA_MACAS
37 72 80.0 141 1 HBA_MACAS
38 72 80.0 141 1 HBA_MACAS
39 72 80.0 141 1 HBA_MACAS
40 72 80.0 141 1 HBA_MACAS
41 72 80.0 141 1 HBA_MACAS
42 72 80.0 141 1 HBA_MACAS
43 72 80.0 141 1 HBA_MACAS
44 72 80.0 141 1 HBA_MACAS
45 72 80.0 141 1 HBA_MACAS

RESULT 1
HBA_TRAFT
ID HBA_TRAFT STANDARD: PRT: 140 AA.

AC P04237: 20-MAR-1997 (rel. 04, created)

DT 20-MAR-1997 (rel. 04, last sequence update)

PT 01-JUN-1994 (rel. 29, last annotation update)

DE HEMOGLOBIN ALPHA CHAIN.

OS Tagelaphus strepsiceros (Greater kudu).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Tragelaphus.

OX NCBI_TaxID=9446;

RN [1]

FP SEQUENCE.

PV MFE438-4527265; P04237-407493;

FA Fekedjeld K., Wiesner H., Braunitzer G.;

RT "Primary structure of the hemoglobin from the greater kudu antelope

PT (Tragelaphus strepsiceros).";

RL Biol. Chem. Heppes Seyler 266:395-402(1985).

DR FTR: A03293; HMDKA.

DR HSSP: P01966; 1HDA.

DR InterPro: IPR000971;

DR InterPro: IPR002318;

DR Pfam: PF00042; globin; 1.

DR PRINTS: PF00042; ALPHAHEM.

DR PROSITE: PS01033; GLOBIN; 1.

KN Heme; Oxygen transport; Respiratory protein; Erythrocyte.

FT METAL 57 57 IRON (HEME DISTAL SITE);

FT METAL 86 86 IRON (HEME PROXIMAL SITE);

SO SEQUENCE 140 AA; 1449 MM; AAD1FA85F4DE24 CRO64;

Query Match 80.0% Score 72; DB 1; Length 140;

Best local Similarity 100.0% Pct Id 94.06;

Motifpos 13; Conservative 0; Mismatchpos 0; Models 0; Gaps 0;

QY 2 FPHFDLSHSAQVC 14

Db 42 FPHFDLSHSAQVC 54

RESULT 2

HBA_TRAFT STANDARD: PRT: 141 AA.

ID HBA_TRAFT

AC P11755: 01-OCT-1999 (rel. 12, created)

DT 01-OCT-1999 (rel. 12, last sequence update)

PT 01-JUN-1994 (rel. 29, last annotation update)

DE HEMOGLOBIN ALPHA-1 CHAIN.

OS Tagelaphus strepsiceros (Greater kudu).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Tragelaphus.

OX NCBI_TaxID=9446;

RN [1]

FP SEQUENCE.

PV MFE438-4527265; P04237-407493;

FA Fekedjeld K., Wiesner H., Braunitzer G.;

RT "Primary structure of the hemoglobin from the greater kudu antelope

PT (Tragelaphus strepsiceros).";

RL Biol. Chem. Heppes Seyler 266:395-402(1985).

DR FTR: A03293; HMDKA.

DR HSSP: P01966; 1HDA.

DR InterPro: IPR000971;

DR InterPro: IPR002318;

DR Pfam: PF00042; globin; 1.

DR PRINTS: PF00042; ALPHAHEM.

DR PROSITE: PS01033; GLOBIN; 1.

KN Heme; Oxygen transport; Respiratory protein; Erythrocyte.

FT METAL 57 57 IRON (HEME DISTAL SITE);

FT METAL 86 86 IRON (HEME PROXIMAL SITE);

SO SEQUENCE 140 AA; 1449 MM; AAD1FA85F4DE24 CRO64;


```

CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DE PIR: A02295; HAEMK.
DR HSSP: P01966; 1FDA.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00342; globin_1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1.
KW Heme: Oxygen transport; Respiratory protein; Erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA: 15115 MW: 85AD213393E1FPA4 C7664.

Query Match 80.0% Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred No. 9 9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 2 FPHFDLSHSAQV 14
DB 43 FPHFDLSHSAQV 55

RESULT 6
HBA_ANTPA STANDARD: PRT: 141 AA
AC P14387;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS Antrous pallidus (pallid bat).
OC Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Placentalia;
OC Mammalia; Eutheria; Chordata; Mammalia; Eutheria; Placentalia;
OC Antrous.
OC Antrous.
OC NCBI_TaxID=9440;
RN 11
RP SEQUENCE.
KA MEDLINE=86050104; PubMed=3675871;
KA Kiehlenschmid T., Koop R.F., Brannstrom G.;
RT "The primary structure of the pallid bat (Antrous pallidus,
RT Chiroptera) hemoglobin."
RL Biol. Chem. Hoppe-Seyler 368:1197-1202(1987).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DE PIR: A02292; A02702.
DR HSSP: P01922; 1FDH.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1.
KW Heme: Oxygen transport; Respiratory protein; Erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA: 15107 MW: F6625666547646 C7664.

Query Match 80.0% Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred No. 9 9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 2 FPHFDLSHSAQV 14
DB 43 FPHFDLSHSAQV 55

RESULT 7

```

```

HBA_ATEGE STANDARD: PRT: 141 AA.
ID HBA_ATEGE
AC P01927;
DE 21-JUL-1986 (Rel. 01, Created)
DE 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA CHAIN.
OS Ateles geoffroyi (Black-headed spider monkey), and
OS Saguinus oedipus (Cotton top tamarin).
OC Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Placentalia;
OC Mammalia; Eutheria; Chordata; Mammalia; Eutheria; Placentalia;
OC NCBI_TaxID=9509; 9490;
RN 11
RP SEQUENCE.
KW SPECIES=A.geoffroyi;
KA MEDLINE=75059126; PubMed=4215720;
KA Matsuda G., Maita T., Sanyama Y., Setoyuchi M., Ota Y., Araya A.,
KA Goodman M., Barnabas J., Prychodko W.;
RT "Studies on the primary structures of alpha and beta polypeptide
RT chains of adult hemoglobin of the spider monkey (Ateles geoffroyi)."
RL Hoppe Seyler's Z. Physik. Chem. 594:1517-1520(1978)
RN 121
RP SEQUENCE.
KW SPECIES=S.oedipus;
KA MEDLINE=8412383; PubMed=6427302;
KA Maita T., Hayashida M., Matsuda G.;
RT "Primary structures of adult hemoglobins of silver marmoset,
RT Callithrix jacchus, and cotton-headed tamarin, Saguinus oedipus."
RL J. Biochem. 95:805-813(1984).
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DE PIR: A02253; HAEMK.
DE PIR: A02253; HAEMK.
DE PIR: B28865; B28865.
DR HSSP: P01922; 1FDH.
DR InterPro: IPR000971; -.
DR InterPro: IPR002338; -.
DR Pfam: PF00042; globin_1.
DR PRINTS: PR00612; ALPHAHAM.
DR PROSITE: PS01033; GLOBIN.1.
KW Heme: Oxygen transport; Respiratory protein; Erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA: 15094 MW: 3B7025A3DE897C2 C7664.

Query Match 80.0% Score 72; DB 1; Length 141;
Best Local Similarity 100.0%; Pred No. 9 9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 2 FPHFDLSHSAQV 14
DB 43 FPHFDLSHSAQV 55

HBA_BISBO STANDARD: PRT: 141 AA.
ID HBA_BISBO
AC P09423;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEMOGLOBIN ALPHA-1 AND ALPHA-11 CHAINS.
OS Bison bonasus (European bison).
OC Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Placentalia;
OC Mammalia; Eutheria; Chordata; Mammalia; Eutheria; Placentalia;
OC NCBI_TaxID=9502;
RN 11
RP SEQUENCE.
KA MEDLINE=86296173; PubMed=3741621;

```


DI 01-NOV-1990 (Rel. 16, Last Created)

DI 01-NOV-1990 (Rel. 16, Last sequence update)

DI 01-OCT-2000 (Rel. 40, Last annotation update)

DE HEMOGLOBIN ALPHA CHAIN.

OS Callithrix jacchus (Shaggy-tailed marmoset).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Placentalia; Callitrichidae;

OC Callithrix.

OC NCBI_TaxId=9482;

OX NCBI_TaxId=9482;

RN [1]

RP SEQUENCE.

RA MEDLINE=8421383; PubMed=647702;

RA Matsuda T., Hayashida M., Matsuda G.;

RT "Primary structures of adult hemoglobins of silvery marmoset,

RT Callithrix jacchus, and cotton-headed tamarin, Saguinus oedipus."

RT J. Biochem. 95:805-813(1984).

CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE

CC VARIOUS PERIPHERAL TISSUES.

CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.

CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.

CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY

DR PIR: S06512; HACUR.

DR HSP: P01922; IPDH.

DR InterPro: IPR000971; -

DR InterPro: IPR002348; -

DR Pfam: PF00042; globin.1.

DR PRINTS: PR00612; ALPHAHAM.

DR PROSITE: PS01033; GLOBIN.1.

KW Heme: Oxygen transport, respiratory protein, erythrocyte,

FT METAL 58 58 IRON (HEME DISTAL LIGAND).

FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).

SQ SEQUENCE 141 AA; 15138 MW; 26B1307E2A68602 CRC64;

QY 2 FPRFDSHOSAOV 14

Db 43 FPRFDSHOSAOV 55

Query Match 80.0%; Score 72; DB 1; Length 141;

Best Local Similarity 100.0%; Pred No. 9; Gap: 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12

HBA_CAMDR STANDARD; PRI: 141 AA.

AC P01974;

DI 21-JUL-1986 (Rel. 01, Created)

DI 21-JUL-1986 (Rel. 01, Last sequence update)

DI 01-OCT-2000 (Rel. 40, Last annotation update)

DE HEMOGLOBIN ALPHA CHAIN.

OS Camelus dromedarius (Dromedary) (Arabian camel), and

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Artiodactyla; Tylopoda; Camelidae; Camelus

OC NCBI_TaxId=9838; 9837;

RN [1]

RP SEQUENCE.

RA MEDLINE=80114125; PubMed=527943;

RA Braunlitzer G., Schrank B., Stangl A., Wiesner H.;

RT "Sequence of hemoglobins from quinea pig and dromedary;"

RT Hofer-Seyler's Z. Physiol. Chem. 360:1941-1946(1979).

RL [2]

RP SEQUENCE.

RC SPECIES=C.dromedarius;

RA Braunlitzer G., Schrank B., Stangl A., Wiesner H.;

RT "Sequence of hemoglobins of the hamster (Mesocricetus auratus) and

RT J. Chem. Soc. Pak. 2:1-7(1980).

CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE

CC VARIOUS PERIPHERAL TISSUES.

CC -1- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.

CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.

CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.

DR PIR: B92758; HACMA.

DR PIR: B91685; HACMA.

DR HSP: P01955; 2ECH.

DR InterPro: IPR000971; -

DR InterPro: IPR002348; -

DR Pfam: PF00042; globin.1.

DR PRINTS: PR00612; ALPHAHAM.

DR PROSITE: PS01033; GLOBIN.1.

KW Heme: Oxygen transport, respiratory protein, erythrocyte,

FT METAL 58 58 IRON (HEME DISTAL LIGAND).

FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).

SQ SEQUENCE 141 AA; 15172 MW; 269D0CF68676E5C CRC64;

QY 2 FPRFDSHOSAOV 14

Db 43 FPRFDSHOSAOV 55

Query Match 80.0%; Score 72; DB 1; Length 141;

Best Local Similarity 100.0%; Pred No. 9; Gap: 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13

HBA_CAMDR STANDARD; PRI: 141 AA.

AC P01970;

DI 21-JUL-1986 (Rel. 01, Created)

DI 21-JUL-1986 (Rel. 01, Last sequence update)

DI 30-MAY-2000 (Rel. 39, Last annotation update)

DE HEMOGLOBIN ALPHA 1 AND ALPHA 2 CHAINS.

OS Capra hircus (goat), Arctocapra lervia (Barbary sheep) (Mouton), and

OS Capra aegagrus (sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Artiodactyla; Ruminantia; Bovidae; Bovidae;

OC Bovidae; Caprinae; Capra.

OC NCBI_TaxId=9925; 9999; 9940;

RN [1]

RP SEQUENCE FROM N.A. (ALPHA 1 AND -2).

RC SPECIES=C.hircus;

RA MEDLINE=82214818; PubMed=6294825;

RA Schon E.A., Wetli S.M., Linquell J.B.;

RT "Gene conversion of two functional goat alpha-globin genes preserves

RT only minimal flanking sequences."

RT J. Biol. Chem. 257:6845-6845(1982).

RN [2]

RP PARTIAL SEQUENCE.

RC SPECIES=C.hircus;

RA MEDLINE=6843123; PubMed=5658544;

RA Huisman T.H.J., Brandt G., Wilson J.B.;

RT "The structure of goat hemoglobins. II. Structural studies of the

RT alpha chains of the hemoglobins A and B."

RT J. Biol. Chem. 243:3675-3686(1968).

RL [3]

RP SEQUENCE.

RC SPECIES=A.lervia;

RA MEDLINE=7016388; PubMed=5449729;

RA Wilson J.B., Wrightstone K.N., Huisman T.H.J.;

RT "Hemoglobin alpha chain duplication in Barbary sheep, Arctocapra

RT lervia, Fallas 1977."

RT Nature 226:354-355(1970).

RL [4]

RP SEQUENCE.

RC SPECIES=Ovis montanus;

RA MEDLINE=67209244; PubMed=6033754;

RA Reale D.;

RT "A partial amino acid sequence for sheep haemoglobin A."

RT Biochem. J. 103:129-140(1967).

RN [5]

Query Match: 83.0%; Score 72; DB 1; Length 141;
 Best local similarity: 100.0%; Pval No; Gap: 0;
 Matches: 3; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 UY 2 PPHFTHSHGSAOV 14
 |||||
 DB 43 PPHFTHSHGSAOV 55

Search completed: June 21, 2001, 10:29:20
 Job time: 24h 50m



GenCore version 4.5
Copyright (c) 1993 - 2000 Compaq Inc.

OM protein - Protein search, using sw model

Run on: June 21, 2001, 10:29:04 ; Search time 33.64 Seconds

(without alignments)
58,995 Million cell updates/sec

Title: US-08-832-443b-2

Perfect score: 90

Sequence: 1 FPFHDSHSAQVC 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 13205027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SPTREMBL_16:
1: sp-archaea:
2: sp-bacteria:
3: sp-fungi:
4: sp-human:
5: sp-invertebrate:
6: sp-mammal:
7: sp-misc:
8: sp-ornithine:
9: sp-phage:
10: sp-plant:
11: sp-rodent:
12: sp-unclassified:
13: sp-vertebrate:
14: sp-virus:

Prod. No. is the number of results produced by chance. To have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	72	80.0	141	6	Q9TS35		Q9TS35 hylobates 1
2	72	80.0	141	6	Q9TS34		Q9TS34 hylobates 1
3	72	80.0	142	4	Q9MYR7		Q9MYR7 homo sapien
4	72	80.0	142	4	Q28744		Q28744 civis arles
5	72	80.0	142	6	Q28744		Q28744 civis arles
6	72	80.0	142	6	Q28745		Q28745 equis cabal
7	72	80.0	142	6	Q28383		Q28383 equis cabal
8	72	80.0	142	6	Q9XSN3		Q9XSN3 equus burch
9	72	80.0	142	6	Q9XSN4		Q9XSN4 equus arcyv
10	72	80.0	142	6	Q9XSK1		Q9XSK1 bubalus bub
11	72	80.0	142	6	Q9XSK9		Q9XSK9 equus cabal
12	72	80.0	142	6	Q9XSK8		Q9XSK8 equus hemo
13	72	80.0	142	6	Q9TVA4		Q9TVA4 equus hemo
14	72	80.0	142	6	Q9TVA3		Q9TVA3 equus burch
15	72	80.0	142	6	Q9TTS1		Q9TTS1 bos taurus
16	72	80.0	142	6	Q9TTS0		Q9TTS0 bos taurus
17	72	80.0	142	6	Q9TTS9		Q9TTS9 bubalus bub
18	72	80.0	142	6	Q9TSS9		Q9TSS9 bubalus bub
19	72	80.0	142	6	Q9TSS8		Q9TSS8 bubalus bub

20	72	80.0	142	6	Q9TSN7		Q9TSN7 bubalus bub
21	72	80.0	142	6	Q9TQNR		Q9TQNR equus asinu
22	71	78.9	100	13	Q9W008		Q9W008 phasiatus c
23	71	78.9	100	13	Q9W008		Q9W008 gallus gall
24	71	78.9	146	13	Q91249		Q91249 columba sp.
25	71	78.9	142	13	Q12986		Q12986 columba liv
26	69	76.7	58	11	Q61649		Q61649 mus musculu
27	69	76.7	122	11	Q90K03		Q90K03 mus musculu
28	69	76.7	142	11	Q61287		Q61287 mus musculu
29	61	67.8	140	13	Q12985		Q12985 columba liv
30	54	60.0	141	13	Q90P25		Q90P25 aechelone
31	54	60.0	142	13	Q91YV9		Q91YV9 hylobates
32	53	58.9	99	11	Q61654		Q61654 mus musculu
33	50	55.6	458	4	Q9XV64		Q9XV64 homo sapien
34	50	55.6	648	4	Q9P209		Q9P209 homo sapien
35	49	54.4	141	13	Q9P819		Q9P819 naja naja
36	48	53.3	142	11	Q63910		Q63910 rattus norv
37	44	48.9	556	10	Q9M172		Q9M172 arabispsis
38	43	47.8	172	5	Q25347		Q25347 leishmania
39	43	47.8	635	4	Q18476		Q18476 homo sapien
40	42	46.7	536	4	Q91Y02		Q91Y02 caenorhabdi
41	42	46.7	838	5	P90882		P90882 caenorhabdi
42	42	46.7	1101	10	Q92907		Q92907 cyrtus saty
43	42	46.7	2148	11	Q93505		Q93505 rattus norv
44	41.5	46.1	581	10	Q95X49		Q95X49 arabispsis
45	41.5	46.1	1376	4	P82279		P82279 homo sapien

ALIGNMENTS

RESULT 1
ID Q9TS35 PRELIMINARY: PRT: 141 AA.

AC Q9TS35;
DT 01-MAY-2000 (TREMUR1, 13, created)
DI 01-MAY-2000 (TREMUR1, 13, last sequence update)
DI 01-MAR-2001 (TREMUR1, 16, last annotation update)
DE ALPHA 1 GLOBIN.
OS Hylobates lar (Common gibbon).
OC Mammalia: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Hylobatidae: Hylobates:
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=92406743; PubMed=1526980;
FA Bailey A.D., Stanhope M., Slidom J.L., Goodman M., Shen C.C.,
RT "Randomly duplicated alpha globin genes of gibbon."
RL J. Biol. Chem. 267:18398-18406(1992).
CC -1- SIMILARITY: TO GLOBIN FAMILY.
CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.
DE HESIT, P01922, 1820.
DE InterPro: IPR000971;
DE InterPro: IPR002338;
DE Pfam: PF00642; globin; 1.
DE PRINTS: PR00612; ALPHABEM.
DE PROSITE: PS01933, GLOBIN, 1.
KW Erythrocyte, Heme, Oxygen transport, Respiratory protein,
SEQUENCE 141 AA, 15136 MW, 29136184.67706 kDa;
SV

Query Match: 80.0%, Score 72, 28 6; Length 141;
Best local similarity: 100.0%, Pctd No. 2, 40-05;
Matches: 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPFHDSHSAQVC 14
DB 43 FPFHDSHSAQVC 55

RESULT 2
Q9TS34

every Match	60.0%	Sticks 2.2	1st 4	Length 14.2
lost Local Significance	100.0%	Prod. No. 2.4e-05		
Matches	14	Mismatch	0	Endless
			0	Capas
			0	0

1 - SIMILARITY: 10

1- SIMILARITY: TO GLOBIN FAMILY.
 1- SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.
 EMBL: X70223; (AA49750.1).

DR HSPB: P01966; 1HBA.
 DR InterPro: IPR000971; ...
 DR InterPro: IPR002388; ...
 DR Pfam: PF00042; Globin_1.
 DR PRINTS: PR00612; ALPHAHBM.
 DR PROSITE: PS01033; Globin_1.
 DR Erythrocyte; Heme; Oxygen transport; Respiratory protein.
 SO SEQUENCE 142 AA; 15220 MW; 93AB5B065096D1; CR664.

Query Match 80.0%; Score 72; DB 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHDLSHGSAOV 14
 ID 028383
 AC 028383
 DT 01-MAR-2001 (TREMBL) 16; Last annotation update)
 DE ALPHA GLOBIN CHAIN.
 OS Ovis aries (Sheep).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Kunitaria; Pecora; Bovidae;
 CC Bovidae; Caprinae; Ovis.
 NX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLUOD;
 FX MEDLINE=9528794; PubMed 776911;
 RA Rietveld M S., Casula S., Rando A., Vestri P.
 RT "Sheep alpha-globin gene sequences: implications for their concerted
 evolution and for the designation of the 3' genes".
 RL J. Mol. Evol. 40:449-453(1995).
 CC -1- SIMILARITY: TO GLOBIN FAMILY.
 CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.
 DR EMBL: X79214; AAA4751.1;
 DR HSPB: P01966; 1HBA.
 DR InterPro: IPR000971; ...
 DR InterPro: IPR002388; ...
 DR Pfam: P01042; Globin_1.
 DR PRINTS: PR00612; ALPHAHBM.
 DR PROSITE: PS01033; Globin_1.
 DR Erythrocyte; Heme; Oxygen transport; Respiratory protein.
 SO SEQUENCE 142 AA; 15178 MW; 070P4DPF887033; CR664;

Query Match 80.0%; Score 72; DB 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHDLSHGSAOV 14
 ID 028383
 AC 028383
 DT 01-NOV-1996 (TREMBL) 01; Created)
 DT 01-NOV-1996 (TREMBL) 01; Last sequence update)
 DT 01-MAR-2001 (TREMBL) 16; Last annotation update)
 DE B11 ALHA-2 CHAIN.
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

QY 2 PFHDLSHGSAOV 14
 ID 028383
 AC 028383
 DT 01-NOV-1996 (TREMBL) 01; Created)
 DT 01-NOV-1996 (TREMBL) 01; Last sequence update)
 DT 01-MAR-2001 (TREMBL) 16; Last annotation update)
 DE B11 ALHA-2 CHAIN.
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

Query Match 80.0%; Score 72; DB 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98215145; PubMed 2815578;
 RA Clegg J.B.;
 RT "Gene conversions in the horse alpha globin gene complex".
 PL Mol. Biol. Evol. 4:492-503(1987).
 FN [2]
 RP SEQUENCE FROM N.A.
 RA Clegg J.B.;
 RT "Schwartz J. (1994) The PHE database from databases.
 CC -1- SIMILARITY: TO GLOBIN FAMILY.
 CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.
 DR EMBL: M17901; AAA30947.1;
 DR HSPB: P01959; 1HBE.
 DR InterPro: IPR000971; ...
 DR InterPro: IPR002388; ...
 DR Pfam: PF00042; Globin_1.
 DR PRINTS: PR00612; ALPHAHBM.
 DR PROSITE: PS01033; Globin_1.
 DR Erythrocyte; Heme; Oxygen transport; Respiratory protein.
 SO SEQUENCE 142 AA; 15229 MW; 93AB5B065096D1; CR664;

Query Match 80.0%; Score 72; DB 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHDLSHGSAOV 14
 ID 028383
 AC 028383
 DT 01-NOV-1996 (TREMBL) 12; Created)
 DT 01-NOV-1996 (TREMBL) 12; Last sequence update)
 DT 01-MAR-2001 (TREMBL) 16; Last annotation update)
 DE ALPHA 1 HEMOGLOBIN.
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 NX NCBI_TaxID=9790;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99055770; PubMed 9847419;
 RA Oakenfull E.A., Clegg J.B.;
 RT "Phylogenetic relationships within the equus genus and the evolution
 of alpha and beta globin genes".
 RL J. Mol. Evol. 47:772-783(1998).
 CC -1- SIMILARITY: TO GLOBIN FAMILY.
 CC -1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.
 DR EMBL: U79200; AAB94461.1;
 DR HSPB: P01958; 1HBE.
 DR InterPro: IPR000971; ...
 DR InterPro: IPR002388; ...
 DR Pfam: PF00042; Globin_1.
 DR PRINTS: PR00612; ALPHAHBM.
 DR PROSITE: PS01033; Globin_1.
 DR Erythrocyte; Heme; Oxygen transport; Respiratory protein.
 SO SEQUENCE 142 AA; 15220 MW; 979F83BD030317E0; CR664;

Query Match 80.0%; Score 72; DB 6; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHDLSHGSAOV 14
 ID 028383
 AC 028383
 DT 01-NOV-1996 (TREMBL) 01; Created)
 DT 01-NOV-1996 (TREMBL) 01; Last sequence update)
 DT 01-MAR-2001 (TREMBL) 16; Last annotation update)
 DE B11 ALHA-2 CHAIN.
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

FX MEMLINF:99065770; PubMed:984719;
 RA Oakenfull E.A., Clegh J.B.:
 "Physico-chemical relationships within the genus *Egus* and the evolution
 of alpha and theta globin genes.";
 RL J. Mol. Evol. 47:772-783(1998).
 RN [2].
 RP SEQUENCE FROM N.A.
 RA Oakenfull E.A., Clegh J.B.:
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO GLOBIN FAMILY.
 CC -1- SIMILARITY: TO ALPHA HEMOGLOBIN FAMILY.
 DE EMBL: AF129139; AAB93466.1;
 DE EMBL: AF129138; AAB93465.1; JOINED.
 DR HSPB: P01936; 118E.
 DR InterPro: IPRO00971; -;
 DR InterPro: IPRO0338; -;
 DR Pfam: PF00042; Globin_1.
 DR PRINTS: PR00612; ALPHAHEM.
 DR PROSITE: PS01033; GLOBIN.1
 KW ERYTHROCYTE; HEMO; OXYGEN; TRANSFER; RESPIRATORY PROTEIN.
 SC SIGNIFIC: 142 AA; 1531 MW; 82KAC21P99049924 GENE41.

Query Match	89.9%	Score 72	1P 67	Length 142
Best Local Similarity	100.0%	Prot No	2 4e-05	
Matches	15, Conservative	0, Mismatches	0, Indels	0, Gaps
QY	2	FFHFDLSHSAOV	14	
DB	44	FFHFDLSHSAOV	56	

RESULT	13	
09TVAA4		
ID	09TVAA4	PRELIMINARY:
AC	09TVAA4	PRT: 142 AA.
DT	01-MAY-2000 (TRFEMRel 13, Created)	
DT	01-MAY-2000 (TRFEMRel 13, Last sequence update)	
DT	01-MAR-2001 (TRFEMRel 16, Last annotation update)	
DE	ALPHA 1 HEMOGLOBIN.	
OS	Equus hemionus (Kulan).	
OC	Eukaryota, Metazoa: Chordata, Craniata, Vertebrata, Euteleostomi:	
OC	Mammalia, Eutheria: Perissodactyla: Equidae: Equus.	
CX	NCBI_TaxID=8794;	
FN	[1]	
RP	SEQUENCE FROM N.A.	
FX	MEF01NF 6066770: P00MD9-9847419;	
RA	Oakrent, F.A., Glegg J.B.:	
KT	"Phylogenetic relationships within the genus Equus and the evolution	
RT	of alpha and theta globin genes."	
RL	J. Mol. Evol. 47:772-783(1998).	
	[2]	
RP	SEQUENCE FROM N.A.	
RA	Oakrent, F.A., Glegg J.B.:	
RL	Submitted (FEB-1999) to the EMBL/GenBank/TrEMBL databases.	
CC	-1- SIMILARITY: TO GLOBIN FAMILY.	
CC	-1- SIMILARITY: TO ALPHA HAEMOGLOBIN FAMILY.	
DR	EMBL: A0729137; AAB93466.1;	
DR	EMBL: A0729136; AAB93465.1; J01NFD	
DR	HSSP: P61958; 11HE.	
DR	InterPro: IPR000971; -;	
DR	InterPro: IPR002388; -;	
DR	Pfam: PF00042; globin.1.	
DR	PRINTS: PRC0612; ALPHAHHEM.	
DR	PROSITE: PS01033; GLOBIN.1	
KM	Erythrocyte: Hemo: Oxygen Transport: Respiratory Protein.	
SQ	SEQUENCE: 142 AA: 15306 MW: 8280376;14310192 Gb004;	

Query Match	80.0%;	Score 72;	DB 6;	Length 142;
Best Local Similarity	100.0%;	Pred. No. 2.4e-05;		
Matches	18;	Conserved*120	0;	Mismatch*400
			0;	Indels
			0;	Gaps

07	2	PPHPII.SHGSAQV	14
1b	44	PPHPII.SHGSAQV	56

RESULT	14		
QTYV3			
10	QTYV3	PRELIMINARY:	PRG: 14Z AA.
AC	QTYV3		
DT	01-MAY-2000 (TEMBELRE, 13, Created)		
DT	01-MAY-2000 (TEMBELRE, 13, last sequence update)		
DT	01-MAY-2001 (TEMBELRE, 15, last annotation update)		
DE	ALPHA 2 HEMOGLOBIN.		
OS	Equus burchelli (Platus zebra) (Equus quagga).		
GC	Eumeryda: Melzoda: Chordata: Craniata: Vertebrata: Euteleostomi:		
GC	Mammalia: Euliparia: Perissodactyla: Equidae: Equus.		
OX	NCBI-TaxID=9790.		
RN	111		
RP	SEQUENCE FROM N.A.		
EX	MDL:MF 99065770; PDBID 9947419;		
EA	DAKENTU E.A., Olaga J.B.;		
RT	"phylogenetic relationships within the genus Equus and the evolution		
EL	of alpha and theta globin genes."		
EL	J. Mol. Evol. 47:772-783(1998).		
RP	121		
RP	SEQUENCE FROM N.A.		
RA	DAKENTU E.A., Olaga J.B.;		
PL	Submitted (FBI-1990) to the FBI/Chaback/DHRT databases.		
CC	-1- SIMILARITY: TO GLOBAL FAMILY.		
CC	-1- SIMILARITY: TO ALPHA HAMB-GLOBIN FAMILY.		
DK	EMBL: AF129135; AAR93662.1;		
DK	EMBL: AF129134; AAR93662.1; JOINED.		
DK	HSSP: P01958; 11PF.		
DK	InterPro: IPR000971;		
DK	InterPro: IPR002338;		
DK	PFam: PF00042; globin_1.		
DK	PRINTS: PR00612; ALPHAHAM.		
DK	PROSITE: PS01033; GLOBIN_1.		
DK	Truncat... Homo, oxypne (transpirt, respiratory protein).		
SC	SEQUENCE 14Z AA: 15190 MW: 167837RC00010865: 16794;		

		(89.7%)	Scores 72% (48/6)	Length 142
		Root Local Similarity	100.0%	Pred. No. 2, 4 or 10?
	Matches	13:	Conserved 100	Mismatches 0;
				Totals 0;
				Gaps 0;
Cy	2 FINEDISNOV	14		
FD	44 FINEDISNOV	56		

CC	RT	PT	PEL.MINAV:	PEL:	AA:
CC	15	097TS1			
AC	097TS1				
DT	01-MAY-2000 (TREMURel. 14, Created)				
DT	01-MAY-2000 (TREMURel. 14, Last sequence update)				
DT	01-MAY-2001 (TREMURel. 16, Last annotation update)				
DE	HEROGLONIN ALPHA CHAIN.				
GN	ALPHA GLOBIN.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eumetazoa; Cetartiodactyla; Perissodactyla; Artiodacta;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RN	11;				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=POLLOIAN;				
RA	PULLIO R. PIOTROWSKI F.; VINCENTI D.; CAMPANELLO G.; DI IACCA A.;				
RT	"Nucleotide sequence of Polloian cattle (Bos taurus primigenus) alpha				
RT	globin genes."				
PT	Chab. Prot. (MAY-1999) -- "doi:10.1006/prot.1999.28148" databases.				
CC	15- SIMILARITY TO GLOBIN FAMILY				

9